(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification[7]: C07K 14/315, C12Q 1/68, C12N 15/31, G01N 33/569

(21) International Application Number:
PCT/EP2004/005664

(22) International Filing Date: 26 May 2004 (26.05.2004)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
03450137.9   30 May 2003 (30.05.2003)   EP

(71) Applicant (for all designated States except US): INTER-CELL AG [AT/AT]; Campus Vienna Biocenter 6, A-1030 Wien (AT).

(72) Inventors; and
(75) Inventors/Applicants (for US only): MEINKE, Andreas [DE/AT]; Piettegasse 26/1, A-3013 Pressbaum (AT). NAGY, Eszter [HU/AT]; Taborstrasse 9, A-1020 Wien (AT). HANNER, Markus [AT/AT]; Jacquingasse 5/6, A-1030 Vienna (AT). GELBMANN, Dieter [AT/AT]; Ungergasse 5, A-7163 Andau (AT).

(74) Agent: SONN & PARTNER PATENTANWALTE; Riemergasse 14, A-1010 Wien (AT).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Declaration under Rule 4.17:
— of inventorship (Rule 4.17(iv)) for US only

Published:
— without international search report and to be republished upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: ENTEROCOCCUS ANTIGENS

(57) Abstract: The present invention discloses isolated nucleic acid molecules encoding a hyperimmune serum reactive antigen or a fragment thereof as well as hyperimmune serum reactive antigens or fragments thereof from E.faecalis, methods for isolating such antigens and specific uses thereof.

ENTEROCOCCUS ANTIGENS

The present invention relates to isolated nucleic acid molecules, which encode antigens for *Enterococci, e.g. Enterococci faecalis*, which are suitable for use in preparation of pharmaceutical medicaments for the prevention and treatment of bacterial infections caused by *Enterococci*.

Enterococci are gram-positive bacteria that are normal inhabitants of the alimentary tract of humans and animals. They have been recognized as a cause of infective endocarditis for more than a century {Murray, B., 1990} and have gained prominence over the last two decades as being among the most common pathogens found in hospital-acquired infections, and surgical-site infections {Richards, M. et al., 2000}. The increasing importance of enterococci as nosocomial pathogens can be attributed in part to intrinsic and acquired antibiotic resistance {Murray, B., 1990}; {Rice, L., 2001}. Treatment of multi-drug-resistent enterococcal infections poses a significant challenge to clinicians {Cetinkaya, Y. et al., 2000}; {Gold, H., 2001}, and the potential of these organisms to serve as a reservoir for antibiotic resistance genes is of great concern {French, G., 1998}; {Noble, W. et al., 1992}; {Poyart, C. et al., 1997}.

The classification of enterococci as group D streptococci dates back to the early 1930's. In 1984, enterococci were given formal genus status after serveral studies demonstrated a more distant relationship with the streptococci.

Enterococci are generally considered commensals of the gastrointestinal tract of a variety of organisms including man. Although there are more than 14 different species of enterococci, *E. faecalis* and *E. faecium* are the species most commonly found in humans. Several intrinsic features of *Enterococcus* may allow members of this genus to survive for extended periods of time, leading to its persistence and nosocomial spread. The ability of enterococci to adapt and persist in the presence of detergents may allow them to survive inadequate cleaning regimens, contributing to their persistence in the hospital. The intrinsic ruggedness of enterococci also confers an unusual level of tolerance to several classes of antibiotics including aminoglycosides, -lactams and quinolones. For example, the resistance of enterococci to aminoglycosides results from the ability of enterococci to block the uptake of the drug at the cell wall. Although the mechanism of high-level resistance was determined to be the result of a bifunctional enzyme {Ferretti, J. et al., 1986}, the molecular basis for the intrinsic resistance of enterococci to low-levels of aminoglycosides remains to be determined.

Among the Enterococci, *E. faecium* is unique because it is commonly used in production of fermented foods, and is also used as a probiotic bacterium. In recent years, *E. faecium* has been less acceptable as a food fermentation organism because of concern that this bacterium may be an intermediate host for spreading of antibiotic resistance to bacteria involved in human infections. Despite these concerns, *E. faecium* is still amongst the most common bacteria found in foods fermented by lactic acid bacteria. Many isolates of *E. faecium* have been shown to produce bacteriocins (antimicrobial peptides) that are able to kill or inhibit growth of pathogens such as listeria, clostridia, bacilli, and staphylococci. Such bacteriocins may contribute to the preservative effect of food fermentations, and is one reason why they have been chosen as starter cultures in the production of fermented food. Recently, enterocins have been implemented successfully in treatment of mastitis in cattle.

Besides the applications for food production, as probiotics and in treatment of animal disease, and more importantly, enterococci are emerging opportunistic human pathogens. This is due to their intrinsic pathogenic potential, and, even more because of their ability to rapidly acquire antibiotic resistance genes. *E. faecium* and *E. faecalis* are the causing agents of a large percentage of hospital-acquired infections, including superinfections.

Enterococci normally colonize the gastrointestinal tract of man. They are found in relative abundance in human feces. A close association is likely to exist between enterococci and their host, or the organism would be eliminated due to normal intestinal motility. Many infection-derived enterococcal isolates were found to be clonal, indicating nosocomial transmission. Moreover, a number of studies have documented patient colonization following hospital admission, and have shown that colonization with multiple resistant strains is a predisposing factor for subsequent infection.

One of the enigmas of nosocomial enterococcal infection not easily explained is the ready colonization of an ecology already occupied by members of the same species. As noted, antibiotics lacking substantial anti-enterococcal activity (i.e. antibiotics that do not deleteriously affect indigenous enterococci) are important predisposing factors for infection. These infections are frequently caused by multiple resistant enterococcal isolates that have been exogenously acquired and appear to have out competed indigenous enterococci in the absence of direct selection.

The fact that exogenous, multiple resistant, nosocomially transmitted enterococci efficiently colonize the gastrointestinal tract suggests that they may not compete directly for the same niche as indigenous strains.

Infection caused by the genus *Enterococcus* include a) bacteremia, b) urinary tract infections c) endophthalmitis, d) endocarditis and also wound and intra-abdominal infections. Approximately ¾ of the infections are caused by the species *E. faecalis*, the rest by *E. faecium*.

a) bacteremia
Nosocomial surveillance data for the period October 1986-April 1997 list enterococci as the third most common cause of nosocomial bacteremia, accounting for 12.8% of all isolates. The translocation of enterococci across an intact intestinal epithelial barrier is thought to lead to many bacteremias with no identifiable source. The risk factors for mortality associated with enterococcal bacteremia include severity of illness, patient age, and use of broad spectrum antibiotics, such as third-generation cephalosporins or metronidazole. These studies suggest that high-level aminoglycoside resistance does not affect clinical outcome, and that the presence of the E. faecalis cytolysin (hemolysin) may enhance the severity of the infection.

b) urinary tract infections
Enterococci have been estimated to account for 110,000 urinary tract infections (UTI) annually in the United States. A few studies have been aimed at understanding the interaction of enterococci with uroepithelial tissue. A potential role for the plasmid-encoded aggregation substance in the adhesion of enterococci to renal epithelial cells has been demonstrated. *E. faecalis* harboring the pheromone responsive plasmid pAD1, or various isogenic derivatives, were better able to bind to the cultured pig renal tubular cell line, LLC-PK, than plasmid free cells. Their findings also showed that a synthetic peptide containing the fibronectin motif, Arg-Gly-Asp-Ser, could inhibit binding. This structural motif mediates the interaction between fibronectin and eucaryotic surface receptors of the integrin family.

c) endophthalmitis
Colonization of host tissue may play a role in the pathogenesis of endophthalmitis. Enterococci are among the most destructive agents that cause this post-operative complication of cataract surgery. Experiments designed to determine whether aggregation substance targeted E. faecalis to alternate anatomical structures within the eye showed that enterococci attach to membranous structures in the vitreous, but that such adherence is not dependent on the presence of aggregation substance.

d) endocarditis
Of the diverse infections caused by enterococci, infective endocarditis (IE) is one of the most therapeutically challenging. Enterococci are the third leading cause of infective endocarditis, accounting for 5-20% of cases of native valve IE, and 6-7% of prosthetic valve endocarditis. The presence of the pheromone-responsive plasmid pAD1 enhances vegetation formation in enterococcal endocarditis. Serum from a patient with *E. faecalis* endocarditis was used to identify an antigen selectively expressed in serum but not in broth culture {Lowe, A. et al., 1995}. This protein antigen, designated EfaA, has extensive sequence similarity with several streptococcal adhesions and might function as an important adhesin in endocarditis.

Ampicillin is the therapy of choice for enterococcal infections. For serious enterococcal infection, particularly for endocarditis, aminoglycosides are critical as part of combination therapy with penicillin or ampicillin. Although enterococci are intrinsically resistant to low levels of aminoglycosides, the addition of the cell wall inhibitors to aminoglycoside will result in an enhanced killing by the synergictic

action of the two antimicrobials. With the increasing incidence of high level resistance to aminoglycosides and penicillins, vancomycin has become the only choice available for the treatment of enterococcal infections. Then, vancomycin resistance was reported in clinical isolates of enterococci in 1988, followed by an outbreak caused by vancomycin-resistant enterococci (VRE). In U.S. hospitals the percentage of nosocomial enterococci resistant to vancomycin increased from 0.3% in 1989 to 7.9% in 1993. Among patients in intensive care units with nosocomial infections an increase was even more dramatic; from 0.4% in 1989 to 13.6% in 1993, a 34-fold increase in the 4-year period.

In addition to a higher mortality rate, vancomycin-resistant enterococcal infections cost on average about $25,000 more to treat and doubled the patients' length of stay in the hospital.

The dramatic increase in vancomycin resistance, especially among E. faecium isolates, indicates that enterococcal infection will pose an increasing challenge in the future. An obvious therapeutic alternative is vaccination with the aim to induce protective immune responses, which prevents or attenuates infections.

Vaccine development is hindered by the lack of sufficient knowledge about the elements of protective immunity against enterococcal infections. There are reports that neutrophil mediated killing of enterococci was largely a function of complement with antibody playing a less essential but potentially important role, though additional evidence for the importance of anti-enterococcal antibodies in promoting clearance by opsonophagocytic killing was recently reported (Gaglani, M. et al., 1997).

The importance of surface proteins in human immunity to Enterococcus already has been appreciated. It is apparent that all clinical isolates express surface proteins with activity relevant to host immune defense. The enterococcal surface protein (Esp) {Shankar, V. et al., 1999}, gelatinase, cytolysin {Haas, W. et al., 2002} and aggregation substance (AS) surface protein {Sussmuth, S. et al., 2000} are well-characterized biochemically and genetically, and have also been shown to be immunogenic {Xu, Y. et al., 1997}. In an animal model of infective endocarditis specific antibodies against the aggregation substance were still not protective {McCormick, J. et al., 2001}.

Thus, there remains a need for an effective treatment to prevent or ameliorate enterococcal infections. Vaccines capable of showing cross-protection against the majority of Enterococcus strains causing human infections could also be useful to prevent or ameliorate infections caused by all other enterococcal species, namely E. faecalis and E. faecium.

A vaccine can contain a whole variety of different antigens. Examples of antigens are whole-killed or attenuated organisms, subfractions of these organisms/tissues, proteins, or, in their most simple form, peptides. Antigens can also be recognized by the immune system in form of glycosylated proteins or peptides and may also be or contain polysaccharides or lipids. Short peptides can be used since for example cytotoxic T-cells (CTL) recognize antigens in form of short usually 8-11 amino acids long peptides in conjunction with major histocompatibility complex (MHC). B-cells can recognize linear epitopes as short as 4-5 amino acids, as well as three-dimensional structures (conformational epitopes). In order to obtain sustained, antigen-specific immune responses, adjuvants need to trigger immune cascades that involve all cells of the immune system necessary. Primarily, adjuvants are acting, but are not restricted in their mode of action, on so-called antigen presenting cells (APCs). These cells usually first encounter the antigen(s) followed by presentation of processed or unmodified antigen to immune effector cells. Intermediate cell types may also be involved. Only effector cells with the appropriate specificity are activated in a productive immune response. The adjuvant may also locally retain antigens and co-injected other factors. In addition the adjuvant may act as a chemoattractant for other immune cells or may act locally and/or systemically as a stimulating agent for the immune system.

Currently vaccines against enterococcal infection are only in the research stages of development. Efforts are focused not only on capsular polysaccharide (CPS) as immunogens {Huebner, J. et al., 2000}, but also on virulence factors and membrane/surface proteins.

The development of protein conjugated vaccines are no doubt a great new addition to the amarmatorium

in the battle against enterococcal infections, but the vaccine can contain only a limited number of enterococcal proteins and given adequate ecological pressure, variation of the pathogencity island and plasmids by non-vaccine clinical isolates remains a real threat. Morover polysaccharide antigens used for active immunization do not provide immunological memory in humans. Conjugation of CPS to non-enterococcal related immunogenic protein carriers (e.g. tetanus toxoid, cholera toxin B subunit, etc.) has been shown to beneficial in inducing higher concentrations of antibodies in vaccines, but it does not provide pathogen-specific B cell and T cell epitopes which would recruit memory B and T cells during a real infection to support the most effective host response. To be able to supplement the enterococcal vaccines with proteins fulfilling these criteria it is necessary to identify conserved immunogenic enterococcal-specific surface proteins.

There is a great potential for passive antibody-based therapy. There have been already attempts to use human intravenous immunoglobulin (IVIG) preparations for prevention. Recent advances in the technology of monoclonal antibody production provide the means to generate human antibody reagents and reintroduce antibody therapies, while avoiding the toxicities associated with serum therapy. Immunoglobulins are an extremely versatile class of antimicrobial proteins that can be used to prevent and treat emerging infectious diseases. Antibody therapy has been effective against a variety of diverse microorganisms reviewed in {Burnie, J. et al., 1998}. Anti-enterococcal mAB could be given therapeutically to immunosuppressed patient, due to organ transplantation, cancer, HIV infection and other causes.

Certain proteins or enzymes displayed on the surface of gram-positive organisms significantly contribute to pathogenesis, are involved in the disease process caused by these pathogens. Often, these proteins are involved in direct interactions with host tissues or in conceiling the bacterial surface from the host defense mechanisms {Navarre, W. et al., 1999}. E. faecalis is not an exception in this regard. Several surface proteins are characterized as virulence factors, important for enterococcal pathogenicity reviewed in {Jett, B. et al., 1994}. If antibodies to these proteins could offer better protection to humans then polysaccharides, they could provide the source of a novel, protein-based enterococcal vaccine to be used in conjunction with or in place of the more traditional capsular polysaccharide vaccine. The use of some of the above-described proteins as antigens for a potential vaccine as well as a number of additional candidates resulted mainly from a selection based on easiness of identification or chance of availability. There is a demand to identify relevant antigens for E. faecalis in a more comprehensive way.

The present inventors have developed a method for identification, isolation and production of hyperimmune serum reactive antigens from a specific pathogen, especially from *Staphylococcus aureus* and *Staphylococcus epidermidis* (WO 02/059148). However, given the differences in biological property, pathogenic function and genetic background, *Enterococcus faecalis* is distinctive from *Staphylococcus* strains. In order to identify relevant serum sources three major types of human sera were collected from healthy adults, as well from patients with enterococcal infections and naïve individuals, young children between 5 and 10 months of age, after they already lost maternal antibodies (as negative controls). A large percentage of individuals are exposed to enterococci in the environment that can induce antibodies in the host. Disease, which mainly occurs in hospitals, might be associated with low levels of specific antibodies against Entercocci, and consequently less efficient phagocytic elimination. To select for appropriate screening reagents, a series of immunoassays (mainly ELISA and immunoblotting) were performed with bacterial lysate and culture supernatant proteins to measure anti-*E. faecalis* IgG antibody levels. Sera from high titer individuals were included in the genomic-based antigen identification.

The genomes of the two bacterial species *E. faecalis* and *S. aureus* by itself show a number of important differences. The genome of *E. faecalis* contains app. 3.22 Mb, while *S. aureus* harbours 2.85 Mb. They have an average GC content of 37.5 and 33%, respectively and approximately 1/3 of the encoded genes are not shared between the two pathogens. In addition, the two bacterial species require different growth conditions and media for propagation. A list of the most important diseases, which can be inflicted by the

two pathogens is presented below. *S. aureus* causes mainly nosocomial, opportunistic infections: impetigo, folliculitis, abscesses, boils, infected lacerations, endocarditis, meningitis, septic arthritis, pneumonia, osteomyelitis, scalded skin syndrome (SSS), toxic shock syndrome. *E. faecalis* causes mainly infections which are not highly toxigenic, highly invasive, or highly infectious by most measures. They do, nevertheless, cause a substantial amount of human disease such as bacteremia, urinary tract infections, endocarditis and intra-abdominal infections.

The complete genome sequence of *E. faecalis* V583, a vancomycin-resistent clinical isolate, was determined by the random shotgun sequencing strategy (GenBank accession number for chromosome and the plasmids are as follows: AE016830 (chromosome), AE016833 (pTEF1), AE016831 (pTEF2), AE016832 (pTEF3)); see www.tigr.org/tigrscripts/CMR2/CMRHomePage.spl). {Paulsen, I. et al., 2003}.

The complete genome of *E. faecium* strain DO (ATCC BAA-472, TEX16, TX0016) with an estimated genome size of 2.8 Mbp has been sequenced. The genome is currently being computed and annotated by the Baylor College of Medicine's Human Genome Sequencing Center and the University of Texas Center for the Study for Emerging and Re-emerging Pathogens (CSERP); see: www.hgsc.bcm.tmc.edu/microbial/Efaecium/

The problem underlying the present invention was to provide means for the development of medicaments such as vaccines against *E. faecalis* infection. More particularly, the problem was to provide an efficient, relevant and comprehensive set of nucleic acid molecules or hyperimmune serum reactive antigens from *E. faecalis* that can be used for the manufacture of said medicaments.

Therefore, the present invention provides an isolated nucleic acid molecule encoding a hyperimmune serum reactive antigen or a fragment thereof comprising a nucleic acid sequence which is selected from the group consisting of:
   a) a nucleic acid molecule having at least 70% sequence identity to a nucleic acid molecule selected from Seq ID No 1-2, 4-8, 10, 12-18, 20-23, 25-26, 29-43, 45-62, 64-74, 76-77, 79-83, 85-89, 91-92, 94-114, 117-126, 128-146, 148-170, 373, 375, 379-381, 387, 392, 394, 397-399, 407-408, 410-411 and 415-424.
   b) a nucleic acid molecule which is complementary to the nucleic acid molecule of a),
   c) a nucleic acid molecule comprising at least 15 sequential bases of the nucleic acid molecule of a) or b)
   d) a nucleic acid molecule which anneals under stringent hybridisation conditions to the nucleic acid molecule of a), b), or c)
   e) a nucleic acid molecule which, but for the degeneracy of the genetic code, would hybridise to the nucleic acid molecule defined in a), b), c) or d).

According to a preferred embodiment of the present invention the sequence identity is at least 80%, preferably at least 95%, especially 100%.

Furthermore, the present invention provides an isolated nucleic acid molecule encoding a hyperimmune serum reactive antigen or a fragment thereof comprising a nucleic acid sequence selected from the group consisting of
   a) a nucleic acid molecule having at least 96%, preferably at least 98 %, especially 100 % sequence identity to a nucleic acid molecule selected from Seq ID No 3, 9, 11, 24, 27, 44, 63, 75, 84, 115-116, 127, 374, 376-378, 382-386, 388-391, 393, 395-396, 400-406, 409 and 412-414,
   b) a nucleic acid molecule which is complementary to the nucleic acid molecule of a),
   c) a nucleic acid molecule comprising at least 15 sequential bases of the nucleic acid molecule of a) or b)
   d) a nucleic acid molecule which anneals under stringent hybridisation conditions to the nucleic acid molecule of a), b) or c),

- 6 -

e) a nucleic acid molecule which, but for the degeneracy of the genetic code, would hybridise to the nucleic acid defined in a), b), c) or d).

According to another aspect, the present invention provides an isolated nucleic acid molecule comprising a nucleic acid sequence selected from the group consisting of

a) a nucleic acid molecule selected from Seq ID No 90, 147.

b) a nucleic acid molecule which is complementary to the nucleic acid of a),

c) a nucleic acid molecule which, but for the degeneracy of the genetic code, would hybridise to the nucleic acid defined in a), b), c) or d).

Preferably, the nucleic acid molecule is DNA or RNA.

According to a preferred embodiment of the present invention, the nucleic acid molecule is isolated from a genomic DNA, especially from a E. faecalis genomic DNA.

According to the present invention a vector comprising a nucleic acid molecule according to any of the present invention is provided.

In a preferred embodiment the vector is adapted for recombinant expression of the hyperimmune serum reactive antigens or fragments thereof encoded by the nucleic acid molecule according to the present invention.

The present invention also provides a host cell comprising the vector according to the present invention.

According to another aspect the present invention further provides a hyperimmune serum-reactive antigen comprising an amino acid sequence being encoded by a nucleic acid molecule according to the present invention.

In a preferred embodiment the amino acid sequence (polypeptide) is selected from the group consisting of Seq ID No 171-172, 174-178, 180, 182-188, 190-193, 195-196, 199-213, 215-232, 234-244, 246-247, 249-253, 255-259, 261-262, 264-284, 287-296, 298-316, 318-340, 425, 427, 431-433, 439, 444, 446, 449-451, 459-460, 462-463 and 467-476.

In another preferred embodiment the amino acid sequence (polypeptide) is selected from the group consisting of Seq ID No 173, 179, 181, 194, 197, 214, 233, 245, 254, 285-286, 297, 426, 428-430, 434-438, 440-443, 445, 447-448, 452-458, 461 and 464-466.

In a further preferred embodiment the amino acid sequence (polypeptide) is selected from the group consisting of Seq ID No 260, 317.

According to a further aspect the present invention provides fragments of hyperimmune serum-reactive antigens selected from the group consisting of peptides comprising amino acid sequences of column "predicted immunogenic aa" and "location of identified immunogenic region" of Table 1a and Table 1c; the serum reactive epitopes of Table 2, especially peptides comprising amino acids 4-10, 14-21, 30-36, 59-68, 77-82, 87-93, 96-105, 112-121, 125-133, 135-141, 150-162, 164-183, 192-203, 207-213, 215-226, 228-234, 241-247, 250-285, 302-308 and 135-148 of Seq ID No 171; 15-57, 60-73, 77-101, 108-134, 136-177, 185-201, 203-217, 226-240, 244-254, 272-277, 283-288, 292-343, 354-370, 380-398, 406-437, 439-453, 473-490, 532-538, 584-590, 595-601, 606-612, 664-677, 679-704, 715-724, 731-753, 759-772, 786-794, 814-862 and 657-684 of Seq ID No 172; 4-9, 15-36, 41-47, 54-60, 75-81, 114-120, 131-146, 152-158, 174-182, 194-202, 208-215, 218-226, 255-271, 276-285, 290-295, 302-311, 318-328, 330-344, 352-359, 365-377, 388-395, 398-405, 426-432, 439-449, 455-500, 505-513, 531-537, 542-552, 554-561, 587-595, 606-612, 718-734, 763-771, 775-782, 792-801, 805-812, 822-828, 830-843, 849-863, 876-894, 905-911, 919-926, 935-947, 949-958, 968-979, 1009-1016, 1029-1045, 1047-

1056, 1076-1081, 1092-1106, 1123-1133, 1179-1200, 1202-1211, 1215-1223, 1287-1299, 1301-1306, 398-431 and 1224-1237 of Seq ID No 173; 17-47, 74-80, 90-97, 126-133, 137-148, 167-173, 179-185, 214-223, 250-255, 270-283, 329-338, 342-350, 352-358, 360-367, 372-383, 398-404, 411-421, 426-432, 435-446, 452-462, 472-479, 515-521, 582-592, 611-618, 623-629, 642-659, 666-673, 678-689, 704-725, 732-737, 744-757, 768-789, 824-834, 842-849, 862-868, 877-887, 904-916, 923-928, 941-947, 962-974, 982-992, 1019-1030, 1032-1044, 1046-1052, 1065-1075, 1077-1087, 1108-1121, 1124-1132, 1137-1151, 1170-1182, 1190-1206, 1208-1214, 1227-1233, 1242-1251, 1254-1273, 1282-1298 and 792-825 of Seq ID No 174; 19-31, 39-67, 82-91; 104-110, 113-128, 149-155, 161-181 and 137-155 of Seq ID No 175; 6-18, 54-63, 69-85, 110-127, 142-156, 158-167, 169-211, 238-246, 248-257, 276-311, 339-349, 371-380, 385-391, 394-403, 421-438, 451-456, 483-489 and 449-468 of Seq ID No 176; 5-15, 24-34, 50-56, 61-83, 98-121, 123-136, 149-162, 166-194, 202-215, 221-227, 229-332, 337-360, 367-402, 404-415, 427-433, 444-462, 471-478, 487-498, 511-518, 521-544, 550-563, 568-574, 580-587, 597-607, 610-616, 624-629 and 468-498 of Seq ID No 177; 11-19, 32-49, 57-63, 65-71, 80-89, 91-133, 166-181, 183-191, 201-230, 234-257, 264-291, 297-303, 305-314, 316-335, 337-354, 359-366, 368-374, 383-388, 394-405, 408-442, 446-470, 483-490, 499-505, 513-538, 544-555, 557-563, 568-590, 598-608, 617-623, 627-636, 641-647, 667-685, 687-693, 710-723, 733-739, 742-754, 769-815 and 366-388 of Seq ID No 178; 4-16, 30-35, 42-53, 67-76, 82-87, 101-108, 112-130, 132-138, 147-152, 161-183, 187-208, 218-225, 265-281, 295-303, 305-317, 322-334, 338-357, 360-368, 370-383, 387-394, 400-419, 421-430 and 255-336 of Seq ID No 179; 19-27, 36-47, 59-66, 76-83, 101-112, 118-125, 142-147, 162-180, 185-196, 225-240, 246-263, 286-304, 314-319, 327-333, 353-367 and 194-214 of Seq ID No 180; 14-43, 70-76, 83-89, 111-117, 122-128, 136-145, 163-170, 175-182, 210-219, 246-251, 266-279, 325-331, 338-346, 348-354, 356-363, 368-379, 422-428, 431-441, 450-456, 466-473, 509-515, 532-542, 549-556, 576-586, 605-612, 617-623, 636-653, 660-667, 674-686, 698-719, 726-731, 738-745, 762-783, 818-828, 836-843, 856-862, 871-881, 903-910, 917-922, 935-941, 956-968, 976-986, 1013-1024, 1026-1038, 1059-1069, 1071-1081, 1102-1115, 1118-1126, 1131-1145, 1164-1176, 1187-1200, 1202-1208, 1221-1227, 1236-1245, 1248-1267, 1273-1292, 252-287 and 805- 844 of Seq ID No 181; 4-18, 21-28, 37-43, 56-70, 101-113, 131-140, 142-150, 162-170, 172-184, 193-204, 209-227, 233-238, 246-264 and 93-168 of Seq ID No 182; 14-20, 44-50, 61-70, 77-96, 99-106, 129-142, 168-181, 187-196, 205-221, 225-241, 277-296 and 257-281 of Seq ID No 183; 18-29, 43-54, 64-76, 78-84, 88-103, 125-149, 159-176, 198-218, 230-242, 256-271, 279-285, 287-293, 300-306, 325-331, 344-351, 357-364, 371-397, 400-414, 419-464, 485-515, 517-526, 529-537, 548-553, 573-580, 584-590, 603-620, 639-661, 676-681, 687-700, 716-761, 772-780, 785-790, 795-803, 823-836, 848-853 and 106-134 of Seq ID No 184; 7-13, 19-42, 44-51, 55-75, 87-97, 99-110, 112-118, 129-135, 141-156, 158-178, 213-220, 230-286, 294-308, 323-338, 345-352, 355-365, 370-392, 394-419, 437-446, 454-460, 474-497, 515-526, 528-546, 569-575 and 128-141 of Seq ID No 185; 12-20, 24-33, 45-70, 73-84, 86-94, 103-116, 118-124, 135-142, 163-170, 176-200, 202-224, 226-234, 237-248, 250-262, 265-287, 296-307, 334-341, 347-356, 361-369, 382-396, 405-415, 418-427, 431-439, 443-449, 452-461, 467-474 and 113-146 of Seq ID No 186; 13-38, 44-50, 52-59, 66-72, 83-94, 103-110, 116-124, 131-137, 158-180, 199-204, 218-233, 241-264, 269-317, 326-342, 350-356 and 70-86 of Seq ID No 187; 29-35, 49-59, 63-84, 86-97, 103-111, 113-126, 130-144, 150-158, 174-198, 221-231, 250-264, 266-273, 291-298, 310-318 and 70-90 of Seq ID No 188; 19-25, 28-52, 60-66, 71-76, 131-142, 149-155, 157-178, 181-213, 218-223, 237-242, 250-257, 260-266, 272-279, 282-290, 321-330, 373-385, 393-407, 441-453, 461-475, 509-521, 529-542, 577-589, 597-610, 643-655, 663-677, 703-718, 729-734, 358-464, 495-570 and 604-685 of Seq ID No 189; 4-29, 51-76, 116-136, 158-173, 179-193, 207-215 and 86-111 of Seq ID No 190; 5-23, 45-70, 79-90, 93-107, 114-122, 142-151 and 18-36 of Seq ID No 191; 9-51, 68-120, 133-149, 158-180, 186-206, 211-220, 222-237, 248-293, 296-310, 317-339 and 248-260 of Seq ID No 192; 14-24, 44-63, 69-98, 108-119, 123-136, 155-161, 164-176, 180-193, 203-208, 215-223, 239-247, 274-281, 283-289, 296-304, 306-313, 315-327, 331-341, 343-353, 357-386, 392-405 and 205-246 of Seq ID No 193; 5-13, 16-23, 36-42, 53-63, 70-83, 96-102 and 14-34 of Seq ID No 194; 4-13, 19-35, 49-56, 59-76, 83-107, 121-134, 144-153, 157-164, 166-186, 194-202, 209-216, 231-253, 257-264 and 98-134 of Seq ID No 195; 16-32, 38-47, 58-68, 78-89, 98-114, 117-123, 132-141, 146-156, 164-170, 179-188, 196-212, 219-230, 232-237, 244-263, 265-274, 278-293, 297-303, 306-326, 339-349, 352-359, 362-367, 373-379, 384-394, 396-406, 423-443, 451-461, 465-484, 490-497, 504-511, 523-533, 537-547, 550-556, 558-566, 573-579, 586-593, 598-609, 615-642, 647-665, 671-686, 693-713, 723-728 and 332-378 of Seq ID No 196; 6-21, 34-44, 58-64, 66-74, 79-87, 114-127, 129-143, 154-162, 174-189, 205-214, 241-262, 266-273, 278-297, 319-324, 328-338, 342-351, 390-398, 409-415, 422-435, 458-464, 471-477, 481-486, 506-531, 534-540, 542-550 and 315-389 of Seq ID No 197; 4-28, 39-45, 52-58, 69-82, 93-115, 122-128, 135-140, 146-163, 177-192, 209-215, 221-232, 271-284, 331-337, 341-352, 360-378, 383-390,

392-401, 409-422, 428-435, 462-470, 474-480, 482-496, 531-539, 541-549, 551-560, 562-569, 576-582, 598-618 and 98-127 of Seq ID No 198; 14-27, 33-47, 61-79, 94-104, 119-133 and 36-60 of Seq ID No 199; 11-22, 29-40, 48-62, 68-73, 96-106, 108-118, 125-149 and 102-126 of Seq ID No 200; 4-11, 45-55, 76-83, 86-102, 105-112, 138-144, 147-153 and 20-48 of Seq ID No 201; 12-20, 28-56, 62-68, 72-82, 93-99, 101-107, 120-133, 135-145, 178-186, 208-232, 279-292 and 36-64 of Seq ID No 202; 6-14, 23-48, 65-82, 92-134, 140-181, 188-219, 228-238, 244-253, 255-261 and 124-145 of Seq ID No 203; 11-25, 31-38, 53-59, 62-71, 89-99, 125-133, 151-157, 182-190, 195-203, 208-215, 219-229, 249-262, 267-275, 287-295, 298-316, 318-325, 328-334, 344-353, 357-363, 371-377, 385-391, 396-415, 425-436, 438-457, 471-485, 538-552, 554-561, 606-625, 630-636, 646-653, 669-679, 695-704, 706-715, 722-747, 763-773 and 714-738 of Seq ID No 204; 10-29, 33-45, 50-60, 70-79, 83-95, 118-124, 136-157, 176-184, 192-205, 207-216, 223-234, 240-246, 258-268, 275-283 and 37-56 of Seq ID No 205; 4-24, 27-38, 46-54, 66-72, 81-97, 112-119, 128-137, 152-157, 173-179, 185-214, 219-225, 227-248, 262-284, 286-295, 301-307 and 117-134 of Seq ID No 206; 26-43, 49-56, 60-71, 74-82, 87-98, 110-116, 131-146, 154-164, 169-178, 183-189, 205-214, 241-246, 255-268, 275-292, 305-314, 316-323, 326-340, 346-363, 397-402, 419-429, 440-446, 452-461, 467-475 and 29-66 of Seq ID No 207; 7-16, 21-39, 48-58, 61-78, 82-89, 109-136, 138-150, 152-176, 182-247, 255-261, 267-332, 336-345, 347-358, 362-368, 371-392, 394-404, 407-472, 490-498, 505-513, 527-544, 554-582, 603-611, 614-620, 632-638 and 500-523 of Seq ID No 208; 24-46, 77-83, 90-97, 99-118, 123-166, 168-177, 204-212, 229-239, 248-262, 273-282, 287-293, 300-319, 321-337, 340-352, 357-366, 391-402, 411-428, 442-450, 464-471, 479-489 and 19-40 of Seq ID No 209; 9-23, 25-34, 53-58, 70-86, 90-97, 99-116, 118-128, 131-141, 185-191, 228-233, 237-253, 255-261, 264-271, 273-280, 302-312, 319-349, 351-359, 362-369, 376-383, 387-394, 398-406, 419-434 and 20-31 of Seq ID No 210; 15-22, 37-43, 71-87, 105-115, 121-127, 135-142, 152-158 and 32-52 of Seq ID No 211; 6-12, 18-29, 37-47, 50-58, 65-83, 85-91, 94-99, 108-123, 142-150, 156-163, 183-193, 215-222, 242-249, 252-258, 261-270, 285-308, 318-326 and 1-95 of Seq ID No 212; 9-61, 65-133, 144-155, 166-173, 175-221, 233-276, 278-313, 329-368 and 210-233 of Seq ID No 213; 11-29, 33-39, 46-51, 65-93, 107-113, 134-143, 147-154, 166-177, 181-188, 214-220, 233-243, 263-269 and 112-128 of Seq ID No 214; 8-46, 110-134, 155-167, 174-183, 188-201, 210-230, 253-258, 267-282, 289-299, 312-319, 322-327, 330-337, 365-381, 389-402, 405-411, 419-425, 439-447, 465-472, 489-512, 525-532, 540-554, 577-589, 591-599, 605-614, 616-624, 633-649 and 503-529 of Seq ID No 215; 34-49, 64-70, 90-118, 124-131, 141-152, 159-165 and 112-128 of Seq ID No 216; 5-15, 26-45, 55-72, 80-85, 93-100, 121-133, 142-148, 154-167, 198-205, 209-215, 241-254, 260-265, 271-279 and 244-270 of Seq ID No 217; 4-36, 38-54, 67-83, 122-153, 159-178, 205-212, 232-242, 244-253, 259-268, 281-288, 298-309, 324-331, 334-370, 372-381, 389-401, 403-429, 441-450, 456-462, 465-471, 473-479, 483-504, 508-518, 537-543, 553-565, 578-584, 592-609, 619-625, 658-667, 669-679, 712-719, 722-729, 737-744, 746-752, 758-765 and 180-226 of Seq ID No 218; 6-17, 23-32, 49-56, 61-67, 76-83, 85-103, 105-111, 120-132, 145-171, 175-185, 191-225, 231-246 and 99-128 of Seq ID No 219; 4-24, 28-48, 52-58, 64-79, 87-100, 104-120, 136-152, 159-166 and 150-163 of Seq ID No 220; 15-27, 65-71, 77-99, 104-121, 128-154, 183-216, 223-229, 234-255, 277-287, 296-308 and 77-97 of Seq ID No 221; 8-18, 44-76, 102-109 and 49-57 of Seq ID No 222; 5-14, 28-40, 42-51, 54-60, 77-83, 89-100, 117-124, 146-172, 176-204, 216-231, 237-244, 267-278, 324-334, 342-348, 396-401, 427-433, 438-450, 452-457, 465-471, 473-481, 491-500, 509-515, 523-544, 550-556, 558-569, 589-595, 606-618, 625-632, 640-649, 665-671, 678-688, 691-698, 717-723, 728-734, 781-789, 800-805, 812-821, 833-868, 873-879, 889-905, 929-939, 988-998, 1046-1061, 1073-1079, 1089-1096, 1115-1124, 1132-1140, 1172-1196, 1220-1226, 1231-1249, 1269-1277, 1287-1301, 1307-1330, 1350-1361, 1369-1378, 1387-1412, 1414-1420, 1422-1439, 1484-1491, 1513-1529, 1552-1561, 1576-1583, 1606-1613, 1617-1640, 1647-1654, 1665-1679, 1686-1698, 1709-1727, 1736-1743, 1750-1757, 1771-1790, 1801-1807, 1817-1823, 1831-1842, 1859-1868, 1870-1882, 1884-1891, 1900-1906, 1909-1914, 1929-1935, 1952-1960, 1974-1988, 2002-2011, 2032-2063, 2071-2081, 2116-2124, 2139-2147, 2149-2159, 2163-2190, 2209-2215, 2245-2253, 2282-2287, 2331-2342, 2360-2370, 2379-2393, 2402-2408, 2414-2421, 2423-2430, 2433-2439, 2442-2450, 2472-2478, 2485-2493, 2495-2503, 2506-2512, 2547-2554, 2558-2564, 2615-2625, 2637-2652, 2692-2698, 2700-2706, 2711-2723, 2731-2740, 2748-2753, 2756-2762, 2765-2772, 2781-2798, 2810-2824, 2844-2852, 2885-2899, 2912-2922, 2937-2944, 2947-2970, 2988-2998, 3016-3025, 3032-3037, 3062-3071, 3129-3148, 3156-3161 and 530-607 of Seq ID No 223; 31-36, 57-62, 79-85, 90-96, 99-112, 120-146, 162-185, 193-203, 208-217, 219-226, 239-253, 283-290, 298-304, 306-321, 340-349, 351-361, 365-372, 386-395, 407-438, 473-486, 537-551, 558-568, 576-594, 598-604 and 75-95 of Seq ID No 224; 14-19, 24-30, 34-42, 45-52, 54-64, 66-82, 95-105, 107-118, 126-163, 171-177, 184-201, 210-215, 260-269, 273-279, 288-304, 321-327, 358-364, 370-375, 380-387, 394-404, 407-413, 421-431, 436-451, 465-474, 504-511, 531-552, 578-587, 614-626, 629-636, 638-671, 691-

715, 719-729, 733-745, 752-759, 768-777, 785-792, 794-802, 805-824, 844-854, 867-880, 885-891, 893-902, 907-924, 939-948, 955-964, 966-975, 987-1000, 1012-1017, 1023-1028, 1050-1071, 1083-1098, 1102-1115, 1133-1146, 1170-1183, 1204-1211, 1213-1223, 1262-1311, 1313-1319, 1346-1355, 1366-1371, 1383-1405, 1409-1414 and 776-819 of Seq ID No 225; 12-27, 30-38, 54-61, 64-74, 82-96, 103-110, 117-125, 134-140, 147-158, 185-201, 218-225, 232-253, 265-280, 319-325, 350-362, 366-372, 376-386, 464-483, 485-490, 511-521, 531-537, 542-559, 564-574, 593-609, 613-619, 637-642, 668-677 and 195-214 of Seq ID No 226; 4-21, 59-67, 73-79, 84-91, 141-151, 186-197, 203-214, 222-227, 237-245, 255-260, 281-292, 294-311, 336-344, 346-355, 422-437, 459-466, 484-491 and 77-109 of Seq ID No 227; 10-45, 52-61, 63-70, 74-102, 112-122, 124-132, 164-178, 181-205, 212-240, 246-256 and 226-247 of Seq ID No 228; 38-50, 53-63, 78-87, 89-111, 126-152, 169-176, 179-186, 193-228, 254-267, 275-282, 288-304, 309-318, 325-341, 346-353, 358-367, 384-395, 404-427, 429-435, 456-465, 467-501, 510-521, 523-536, 541-548, 552-560, 563-584, 589-595, 597-620, 625-637, 639-645, 661-666, 712-729, 734-741, 743-750, 775-806, 809-816, 818-840, 842-850 and 693-714 of Seq ID No 229; 5-17, 30-37, 52-75, 77-86, 88-107, 112-135, 151-160, 178-222, 226-246, 263-270, 279-294, 306-314, 327-342, 345-352, 374-381, 389-416, 422-429, 435-449, 453-467, 473-500, 512-522, 524-531, 542-549, 552-560, 565-571, 575-586, 594-600, 613-619, 625-633, 635-641, 647-653, 667-674, 680-699, 711-729, 735-741, 764-775, 781-786, 792-798, 805-813, 817-825, 833-842, 850-855, 860-866, 869-910, 917-930, 949-990 and 533-562 of Seq ID No 230; 7-14, 39-46, 61-74, 83-89, 93-99, 110-121, 136-150, 172-180, 182-200, 207-216, 223-236, 238-251, 265-271, 280-288, 294-309, 320-336, 339-354, 362-377, 383-389, 401-407, 435-441, 446-453, 460-465, 472-487, 499-511, 518-528, 533-540, 557-570, 572-587, 631-637, 643-658, 663-669, 672-678, 681-687, 695-706, 714-728 and 118-139 of Seq ID No 231; 5-19, 24-30, 56-64, 69-79, 93-100, 102-111, 117-123, 125-133, 174-182, 185-199, 205-224, 268-275, 311-336 and 102-125 of Seq ID No 232; 6-35, 39-45, 57-62, 80-85, 92-106, 117-122, 126-171, 214-223, 253-260, 268-273, 285-291, 295-306, 315-320, 325-336, 361-366 and 172-202 of Seq ID No 233; 4-13, 24-37, 45-51, 58-66, 84-92, 112-121, 132-141, 151-171, 175-195, 204-212, 222-240, 262-268, 276-295, 305-336, 338-348, 354-362 and 160-183 of Seq ID No 234; 10-16, 24-35, 41-73, 78-104, 111-121, 124-139, 141-148, 150-164, 196-215, 224-241, 249-282, 299-307, 315-357, 368-378, 393-401 and 345-367 of Seq ID No 235; 4-32, 48-53, 61-67, 84-104, 112-118 and 106-130 of Seq ID No 236; 21-28, 31-36, 65-81, 98-105, 115-121, 123-131, 136-142, 155-161, 177-190 and 201-232 of Seq ID No 237; 4-15, 21-27, 33-39, 42-56, 58-64, 68-82, 84-90, 92-98, 113-122, 146-162, 168-175, 177-189, 191-203, 249-268, 279-285, 287-304, 328-342, 349-358, 371-378, 387-393, 404-413, 419-425, 467-479, 487-498, 513-524, 528-539, 541-565, 572-579, 595-606, 626-635, 637-642 and 612-626 of Seq ID No 238; 7-13, 52-70, 76-82, 97-106, 110-117 and 13-45 of Seq ID No 239; 5-10, 12-48, 59-64, 87-102, 107-128, 131-140, 154-161, 165-171, 173-215 and 54-74 of Seq ID No 240; 4-11, 19-28, 34-40, 74-81, 87-98, 126-147, 163-171, 184-193, 205-213 and 49-124 of Seq ID No 241; 7-14, 23-29, 35-40, 61-67, 99-106, 111-122, 124-133, 135-161, 187-206, 216-229, 236-245, 262-268, 271-280 and 256-273 of Seq ID No 242; 4-13, 17-37, 47-54, 85-99, 105-113, 120-132, 147-166, 180-186, 192-199, 204-216 and 127-144 of Seq ID No 243; 14-27, 29-37, 52-62, 68-76, 89-96, 117-123, 125-131, 137-145, 166-195, 205-212, 214-222, 228-235, 258-264, 271-281, 288-296, 308-324, 332-339, 355-361, 365-371 and 268-293of Seq ID No 244; 4-21, 30-42, 54-60, 78-85, 90-110, 141-147, 160-168, 176-185, 194-206, 218-225, 230-245, 251-261, 287-293, 295-304, 320-326, 334-347, 351-362, 386-402, 413-423, 427-433, 439-453, 456-477, 480-493, 507-513, 526-539, 574-581, 591-598, 600-609, 614-632, 655-665, 685-691, 703-712, 742-747, 757-775, 797-803, 813-819, 823-829, 880-887, 901-906, 930-944, 948-958, 962-968, 971-995, 1002-1009, 1017-1023, 1036-1053, 1069-1081, 1107-1124, 1129-1152, 1178-1195, 1211-1223, 1249-1266, 1271-1288, 1334-1340, 1346-1367, 1-63 and 171-189 of Seq ID No 245; 4-22, 52-63, 70-75, 94-104, 112-125, 133-141, 176-199, 209-216, 244-259, 287-299, 336-352, 366-372, 386-399, 421-436, 444-449, 457-466, 481-487, 506-529, 531-540 and 295-378 of Seq ID No 246; 9-30, 43-49, 58-75, 86-96, 119-131, 138-147, 162-167, 181-201, 208-214 and 16-121 of Seq ID No 247; 4-27, 52-58, 80-90, 92-100, 108-114, 118-143, 169-176, 189-198, 247-261, 281-287, 307-317, 323-329, 352-363, 372-381, 396-411, 413-426, 429-440, 442-450, 456-461, 468-479 and 1-73 of Seq ID No 248; 4-32, 47-52, 57-63, 71-78, 92-104, 126-142, 153-175 and 145-163 of Seq ID No 249; 17-23, 35-41, 51-70, 73-86, 104-125 and 105-129 of Seq ID No 250; 25-32, 41-50, 75-85, 87-103, 115-122, 138-149, 164-171, 188-210, 212-220, 224-234, 256-273, 288-299, 304-310, 330-336, 357-365, 382-390, 399-405, 414-421, 440-446, 454-461, 480-486, 502-514, 518-540, 543-553, 561-567, 572-580, 582-588, 595-630, 633-651, 672-681, 691-709, 760-767, 813-832, 841-848, 852-866, 873-893, 919-925, 927-933, 940-955, 957-978, 984-997, 1000-1010, 1035-1040, 1044-1051, 1058-1064, 1081-1091, 1097-1124, 1129-1138, 1144-1150, 1158-1165, 1170-1180, 909-936 and 1001-1031 of Seq ID No 251; 4-12, 19-26, 31-41, 49-64, 66-86, 101-117, 119-127, 134-142, 152-161, 163-172, 179-188, 209-218, 234-241,

276-291, 294-300, 307-320, 324-341, 346-356, 373-387, 389-397, 410-416, 418-436, 444-454, 460-472, 481-486, 500-507, 511-535, 541-549, 553-559, 579-586, 602-607, 613-620, 628-640, 654-663, 671-678, 681-691, 709-722, 741-754, 766-774, 778-786, 797-803 and 212-226 of Seq ID No 252; 4-10, 15-27, 34-54, 60-73, 79-88, 101-115, 120-136, 154-162, 167-172, 222-240 and 126-195 of Seq ID No 253; 5-16, 18-25, 29-35, 57-63, 86-91, 107-121, 123-131, 170-179, 185-199, 204-226, 250-255, 262-274, 291-296, 325-347 and 1-38 of Seq ID No 254; 7-19, 22-34, 36-42, 48-54, 60-66, 71-76, 104-110, 118-133, 135-145, 158-164, 167-174, 182-193, 196-204, 217-229, 251-290, 293-299, 309-315 and 288-318 of Seq ID No 255; 43-51, 55-61, 66-73, 80-90, 103-127, 133-142, 174-180, 185-196, 203-210, 229-235, 239-251, 258-266, 272-278, 289-314, 316-326, 340-346, 355-361 and 14-27 of Seq ID No 256; 4-25, 27-33, 35-41, 52-74, 76-89, 99-124, 138-144, 146-159, 167-182, 184-191, 193-206, 211-223, 232-240, 249-257, 270-279, 281-287, 293-310, 322-341, 347-356 and 292-322 of Seq ID No 257; 5-13, 28-38, 43-60, 67-72, 98-116, 122-134, 137-151, 167-174, 177-195, 197-216 and 99-195 of Seq ID No 258; 15-33, 35-42, 48-57, 62-68, 73-91, 107-119, 121-153, 173-194, 205-210, 223-228, 234-241, 243-259, 275-298, 308-315, 327-340, 342-370, 376-391, 398-404, 410-419 and 71-122 of Seq ID No 259; 12-39, 43-64, 87-95, 99-105, 114-126, 128-136, 139-147, 212-225 and 107-141 of Seq ID No 260; 6-33, 40-45, 60-75, 79-86, 121-129, 131-137, 161-167, 172-178, 186-195, 203-212, 236-244, 257-264, 278-294, 306-312, 345-358, 368-381, 386-395, 404-410, 412-418 and 198-270 of Seq ID No 261; 18-31, 34-41, 50-56, 69-83, 99-106, 129-141, 147-153, 159-168, 170-178, 190-198, 200-212, 221-232, 237-255, 261-266, 274-292 and 118-216 of Seq ID No 262; 17-47, 61-67, 87-93, 115-121, 126-132, 140-148, 167-173, 179-186, 214-223, 250-255, 264-272, 282-294, 306-318, 338-353, 358-377, 385-401, 414-420, 433-441, 451-457, 470-480, 505-511, 544-550, 571-581, 600-607, 612-618, 631-648, 655-662, 669-681, 693-714, 721-726, 733-740, 757-778, 813-823, 831-838, 851-857, 866-876, 893-905, 912-917, 930-936, 951-963, 971-981, 1008-1019, 1021-1033, 1035-1041, 1054-1064, 1066-1076, 1097-1110, 1113-1121, 1126-1140, 1159-1171, 1182-1195, 1197-1203, 1216-1222, 1231-1240, 1243-1262, 1268-1287 and 738-828 of Seq ID No 263; 19-28, 40-46, 51-57, 68-74, 81-87, 98-108, 111-121 and 20-36 of Seq ID No 264; 4-17, 19-44, 60-69, 80-86, 110-116 and 33-60 of Seq ID No 265; 8-16, 18-28, 42-50, 53-75, 79-86, 94-99, 122-128, 136-142, 149-163, 166-173, 198-212, 254-272, 288-295, 304-318, 324-329, 343-348, 351-364, 367-383, 389-395, 411-420, 427-436 and 11-56 of Seq ID No 266; 19-25 and 6-24 of Seq ID No 267; 6-39, 59-68 and 44-63 of Seq ID No 268; 5-14, 21-28, 38-53 and 29-41 of Seq ID No 269; 4-13, 31-41, 56-65 and 32-56 of Seq ID No 270; 5-12 and 4-21 of Seq ID No 271; 4-18 and 17-32 of Seq ID No 272; 4-10, 23-33 and 14-30 of Seq ID No 273; 26-34, 44-53 and 35-52 of Seq ID No 274; 1-19 of Seq ID No 275; 4-17, 23-30, 32-37 and 6-23 of Seq ID No 276; 5-33, 40-58, 61-66 and 45-66 of Seq ID No 277; 15-41, 61-67 and 41-65 of Seq ID No 278; 4-12, 16-23, 26-37 and 10-29 of Seq ID No 279; 23-39 and 37-55 of Seq ID No 280; 12-20 and 38-55 of Seq ID No 281; 22-37 and 7-22 of Seq ID No 282; 3-14 of Seq ID No 283; 6-16, 43-65, 71-76 and 17-31 of Seq ID No 284; 4-13, 27-39, 42-69 and 17-32 of Seq ID No 285; 4-12, 26-39 and 10-25 of Seq ID No 286; 2-31 of Seq ID No 287; 6-38, 49-62 and 39-55 of Seq ID No 288; 4-10, 24-30 and 2-19 of Seq ID No 289; 12-17, 25-46 and 15-30 of Seq ID No 290; 4-13 and 2-28 of Seq ID No 291; 30-38 and 17-45 of Seq ID No 292 ; 24-33, 55-61 and 31-61 of Seq ID No 293; 4-26, 34-48 and 15-33 of Seq ID No 294; 9-15 and 1-22 of Seq ID No 295; 4-31 and 14-33 of Seq ID No 296; 5-34, 49-55, 64-82 and 69-83 of Seq ID No 297; 33-45 and 21-39 of Seq ID No 298; 7-14, 24-32, 42-65, 79-86 and 50-64 of Seq ID No 299; 13-27, 33-43, 45-62 and 12-37 of Seq ID No 300; 4-15, 17-32 and 10-26 of Seq ID No 301; 4-9, 11-43, 45-75 and 47-69 of Seq ID No 302; 4-18, 22-37 and 17-34 of Seq ID No 303; 4-14 and 5-24 of Seq ID No 304; 7-33, 35-46 and 1-19 of Seq ID No 305; 13-37, 69-75 and 51-69 of Seq ID No 306; 14-24, 26-34, 37-49, 66-78 and 2-25 of Seq ID No 307; 17-46, 52-57, 59-64 and 54-68 of Seq ID No 308; 4-22 and 13-25 of Seq ID No 309; 8-40, 53-63 and 29-50 of Seq ID No 310; 16-28 and 32-40 of Seq ID No 311; 14-20, 22-28, 39-45 and 2-22 of Seq ID No 312; 4-13 and 12-31 of Seq ID No 313; 15-21 and 2-17 of Seq ID No 314; 4-17 and 20-36 of Seq ID No 315; 4-19 and 9-18 of Seq ID No 316; 4-14 and 3-19 of Seq ID No 317; 4-21, 32-40 and 21-39 of Seq ID No 318; 4-13 and 10-27 of Seq ID No 319; 18-31, 39-47, 75-87, 89-98 and 79-99 of Seq ID No 320; 15-21 and 9-24 of Seq ID No 321; 4-14, 18-27, 30-53, 55-64, 68-74, 81-98 and 22-40 of Seq ID No 322; 7-24, 44-51 and 35-60 of Seq ID No 323; 10-47 and 23-37 of Seq ID No 324; 4-10, 12-46 and 7-22 of Seq ID No 325; 20-27 and 1-13 of Seq ID No 326; 6-19, 41-51 and 9-37 of Seq ID No 327; 4-9, 11-17 and 9-23 of Seq ID No 328; 4-17, 23-38, 46-66, 68-85 and 34-46 of Seq ID No 329; 4-18, 34-59, 75-81 and 61-84 of Seq ID No 330; 6-17 and 7-28 of Seq ID No 331; 4-32, 56-61 and 35-52 of Seq ID No 332; 4-14, 27-71, 74-88, 93-110, 115-120, 124-130, 139-154, 161-172 and 146-171 of Seq ID No 333; 4-21 and 3-15 of Seq ID No 334; 12-17 and 9-26 of Seq ID No 335; 10-21, 45-58 and 51-67 of Seq ID No 336; 59-66, 68-84 and 13-42

of Seq ID No 337; 11-16 and 1-16 of Seq ID No 338; 4-19, 23-37 and 10-30 of Seq ID No 339; 19-27, 35-46, 48-66, 82-88, 99-105, 113-119 and 42-59 of Seq ID No 340; 135-147 of Seq ID No 171; 658-682 of Seq ID No 172; 411-427 and 1226-1246 of Seq ID No 173; 794-817 and 801-824 of Seq ID No 174; 468-492 and 474-495 of Seq ID No 177; 366-388 of Seq ID No 178; 266-291, 287-312 and 308-333 of Seq ID No 179; 197-213 and 195-211 of Seq ID No 180; 252-275, 262-285 and 812-830 of Seq ID No 181; 94-112, 97-120 and 104-128 of Seq ID No 182; 257-281 of Seq ID No 183; 106-134 of Seq ID No 184; 70-86 of Seq ID No 187; 358-383, 378-402, 397-421, 499-524, 520-545, 541-566, 622-646, 641-665 and 660-684 of Seq ID No 189; 248-260 of Seq ID No 192; 15-34 of Seq ID No 194; 112-129 of Seq ID No 195; 333-358 and 353-378 of Seq ID No 196; 316-343, 339-366 and 362-389 of Seq ID No 197; 98-123 and 104-126 of Seq ID No 198; 20-43 and 23-48 of Seq ID No 201; 124-145 of Seq ID No 203; 717-738 of Seq ID No 204; 37-56 of Seq ID No 205; 118-134 of Seq ID No 206; 500-522 of Seq ID No 208; 32-47 of Seq ID No 211; 25-51, 47-73 and 69-95 of Seq ID No 212; 503-529 of Seq ID No 215; 112-128 of Seq ID No 216; 181-199 of Seq ID No 218; 109-121 of Seq ID No 219; 150-163 of Seq ID No 220; 77-97 of Seq ID No 221; 564-586 of Seq ID No 223; 75-94 of Seq ID No 224; 776-798, 784-808 and 794-815 of Seq ID No 225; 196-212, 78-100 and 85-107 of Seq ID No 226; 536-553 of Seq ID No 230; 102-125 of Seq ID No 232; 178-198 of Seq ID No 233; 612-626 of Seq ID No 238; 171-187 of Seq ID No 245; 296-320, 315-339, 334-358 and 353-377 of Seq ID No 246; 47-71 of Seq ID No 247; 1-25, 20-45 and 40-64 of Seq ID No 248; 146-161 of Seq ID No 249; 910-935 and 1007-1030 of Seq ID No 251; 212-226 of Seq ID No 252; 126-152, 148-173 and 169-195 of Seq ID No 253; 288-310 and 293-316 of Seq ID No 255; 293-312 of Seq ID No 257; 154-170 of Seq ID No 258; 72-95, 90-112 and 97-121 of Seq ID No 259; 135-150 and 146-163 of Seq ID No 262; 799-827 of Seq ID No 263; 23-43 and 33-53 of Seq ID No 266; 44-62 of Seq ID No 268; 6-22 of Seq ID No 276; 37-54 of Seq ID No 280; 40-54 of Seq ID No 281; 7-21 of Seq ID No 282; 4-11, 16-34, 48-55, 67-77, 87-106 and 153-183 of Seq ID No 425; 22-40, 49-65, 70-91, 95-109, 111-125, 146-207, 209-216, 219-225, 229-244, 251-270, 274-286, 292-309, 316-329, 335-355, 358-370, 376-388, 392-419, 425-430, 435-441, 448-455, 464-478, 486-515 and 437-465 of Seq ID No 426; 5-19, 25-31, 43-48, 60-79, 88-100, 105-129, 148-171, 187-193, 243-263, 316-322, 334-340, 345-351, 369-378, 381-391, 399-404, 474-483, 502-517, 525-530, 558-568, 579-596, 622-627, 631-638, 644-651, 653-660, 674-680, 687-693, 721-728, 743-753, 760-775, 788-795, 806-813, 821-828, 835-842, 847-859, 868-887 and 300-347 of Seq ID No 427; 5-26, 37-44, 89-97, 112-118, 121-128, 138-154, 157-165, 176-181, 188-198, 205-218, 223-243, 247-253, 260-279 and 76-155 of Seq ID No 428; 4-29, 41-46, 49-68, 82-88, 121-147, 158-164, 187-193, 195-208, 229-236, 244-249, 251-263, 269-275, 307-313, 337-343, 348-381, 392-398, 402-408, 432-438, 85-117 and 194-239 of Seq ID No 429; 5-12, 14-22, 28-34, 40-46, 70-79, 84-129, 152-165, 174-182 and 37-109 of Seq ID No 430; 5-16, 18-52, 54-72, 81-86, 118-126, 136-145, 151-157, 168-180, 209-233, 244-270, 295-302, 315-326, 329-337, 345-352, 364-373, 397-402, 408-418, 424-431, 435-443, 472-480, 483-489, 504-510, 519-527, 549-564, 576-599, 605-637, 641-673 and 91-98 of Seq ID No 431; 23-36, 42-52, 133-140, 151-157, 242-247, 267-277, 295-301, 320-328, 333-339, 345-352, 365-371, 397-403, 415-428, 456-465, 481-487, 489-495, 508-516, 518-527, 585-592, 606-614, 631-637, 643-658, 665-670, 723-728, 737-744, 752-759, 787-793, 835-841, 873-885, 918-928, 938-945, 951-966, 978-988, 1015-1020, 1030-1036, 1044-1052, 1058-1069, 1071-1079, 1081-1088, 1113-1119, 1125-1138, 1141-1147, 1164-1170, 1172-1177, 1190-1200, 1214-1220, 1230-1236, 1239-1245, 1262-1268, 1270-1275, 1288-1298, 1312-1318, 1328-1334, 1337-1343, 1360-1366, 1368-1373, 1386-1396, 1410-1416, 1426-1432, 1435-1441, 1458-1464, 1466-1471, 1484-1494, 1508-1514, 1524-1530, 1533-1539, 1556-1562 and 307-340 of Seq ID No 432; 19-25, 35-41, 44-50, 66-72, 74-79, 92-102, 116-122, 132-138, 141-147, 164-170, 172-177, 190-200, 214-220, 230-236, 239-245, 262-268, 270-275, 288-298, 312-318, 328-334, 337-343, 360-366, 368-373, 386-396, 410-416, 426-432, 435-441, 458-464, 466-478, 504-524, 79-148, 177-246, 275-344 and 373-442 of Seq ID No 433; 7-14, 16-23, 33-39, 46-53, 72-79, 92-115, 123-130, 156-175, 179-187, 214-220, 239-246, 266-274, 302-325, 338-354, 360-370, 375-390, 392-401, 421-428, 430-463 and 29-58 of Seq ID No 434; 4-9, 22-39, 58-65, 72-82, 87-92, 99-104, 107-119, 143-166, 171-177, 194-202, 205-213, 220-228, 231-240, 247-263, 309-315, 317-323, 336-343 and 294-320 of Seq ID No 435; 4-10, 12-18, 24-29, 34-43, 50-65, 70-76, 111-117, 129-138, 152-159, 166-171, 184-195, 200-210, 224-236, 241-251, 274-283, 285-296, 313-319, 332-341, 348-355, 378-386, 410-416, 433-445, 475-482, 523-529, 531-540, 584-596, 626-633, 674-680, 682-688, 738-750, 780-787, 828-834, 836-842, 853-862, 882-887, 893-912 and 604-676 of Seq ID No 436; 15-38, 49-57, 60-99, 103-119, 124-194, 200-206, 215-249, 251-291, 307-313, 315-347, 369-378, 383-390, 393-400, 405-411, 423-435, 440-447, 454-460, 470-486, 490-503, 532-539, 542-549, 551-567, 579-592 and 509-583 of Seq ID No 437; 38-44, 47-88, 95-103, 157-172, 235-240, 250-260, 263-276, 294-300, 312-317, 331-337,

369-391, 412-419, 442-448, 453-463, 490-529, 537-555, 571-580, 600-617, 619-627, 642-648, 682-687, 693-700, 716-722, 738-748, 756-763, 779-789, 796-802, 820-828, 833-840, 846-853, 862-872, 880-887, 894-899, 924-937, 957-963, 1006-1012, 1043-1049, 1063-1069, 1076-1097 and 124-147 of Seq ID No 438; 4-28, 31-49, 60-71, 75-102, 104-114, 134-144, 160-184, 250-257, 277-285, 287-294, 330-338, 345-351, 367-374, 381-388, 393-399, 402-407, 420-426, 443-448, 458-464, 411-436 and 454-488 of Seq ID No 439; 20-27, 45-55, 57-64, 66-77, 98-106, 130-137, 155-165, 167-174, 176-187, 194-203, 208-223, 227-238, 245-251, 257-270, 273-278, 287-299, 330-345, 352-358, 363-385, 392-399, 410-417, 437-443, 467-484, 486-492, 495-500, 504-516, 526-536 and 219-270 of Seq ID No 440; 11-22, 24-31, 46-63, 65-71, 73-88, 95-109, 174-181, 183-201, 204-212, 216-222, 228-233, 241-247 and 142-221 of Seq ID No 441; 8-28, 51-59, 67-84, 93-98, 140-152, 154-162, 183-188 and 91-125 of Seq ID No 442; 10-22, 27-61 and 69-100 of Seq ID No 443; 7-15, 18-26, 94-100, 126-131, 152-165, 219-228, 254-263, 274-292, 297-308, 333-340, 342-352, 354-371, 373-379, 403-410, 420-438, 450-456, 463-470, 489-495, 503-512 and 97-173 of Seq ID No 444; 4-21, 37-43, 49-65, 67-74, 76-90, 113-119, 131-141, 155-173, 175-189, 192-199, 207-221, 247-254, 266-276, 317-322, 337-343, 387-393, 408-428, 439-448, 451-460, 469-479, 482-487, 493-501, 517-523, 533-542 and 480-503 of Seq ID No 445; 11-26, 40-46, 78-86, 93-103, 121-126, 132-138, 166-177, 183-196, 203-212, 214-221, 228-263, 304-311, 323-338, 345-351, 357-363, 379-393, 420-434, 442-448, 518-527, 547-553, 581-591, 602-609, 637-645, 665-674, 687-692, 701-708, 730-739, 796-802, 844-857, 882-888, 903-914, 944-950, 976-983, 1027-1033, 1049-1057, 1066-1072, 1085-1092, 1120-1127, 1137-1144, 1153-1158, 1165-1176, 1181-1187, 1221-1230, 1238-1244, 1269-1274 and 605-632 of Seq ID No 446; 6-47, 57-65, 83-95, 109-121, 138-147, 154-164, 167-177, 194-200, 202-212, 227-234, 240-253, 260-267, 283-291, 320-329, 340-347, 356-364, 412-422, 430-436, 441-459, 465-475, 478-486, 498-507 and 59-84 of Seq ID No 447; 10-21, 58-83, 88-97, 120-126 and 21-51 of Seq ID No 448; 5-39, 56-62, 76-88, 90-114, 138-162, 170-195, 202-221, 228-250, 264-270, 304-355, 374-387, 391-416, 462-471, 526-546, 554-561, 574-579, 639-645, 651-660, 674-682, 689-694 and 666-697 of Seq ID No 449; 6-30, 36-42, 143-157, 176-197, 202-209, 216-233, 241-246, 275-287, 292-299, 315-325, 343-350, 375-380, 397-403, 411-420, 422-434, 441-448, 467-474, 477-499, 555-568, 591-597, 601-609, 623-644, 667-688, 692-698, 703-718, 736-747, 757-766, 782-791, 795-801, 832-840, 859-865 and 226-269 of Seq ID No 450; 6-23, 43-51, 61-67, 73-82, 91-97, 123-130, 149-158, 164-175, 228-234, 240-246, 248-255, 262-272, 326-332, 340-347, 365-371, 377-388, 409-419, 425-431, 438-445, 449-457, 464-470, 496-507, 559-568, 575-581, 603-608, 617-623, 633-639, 648-654, 659-670, 695-701, 734-752, 806-814, 816-829, 861-868, 891-899, 904-909, 937-945, 947-960, 978-983, 992-999, 1022-1031, 1068-1076, 1078-1091, 1109-1114, 1123-1130, 1153-1162, 1199-1207, 1209-1222, 1254-1261, 1284-1293, 1330-1338, 1340-1353, 1371-1376, 1385-1392, 1415-1421, 1433-1438, 1460-1465, 1470-1492 and 1422-1458 of Seq ID No 451; 82-94, 111-118, 125-131, 206-212, 261-266, 310-320, 328-338, 345-351, 353-360, 414-420, 424-434, 440-447, 451-500, 506-516, 548-561, 566-572, 584-591, 601-622, 630-636, 650-659, 661-667, 674-699, 703-711, 717-729, 736-744, 752-759, 765-771, 813-822, 826-842, 852-868, 870-877, 881-895, 897-906, 913-922 and 602-671 of Seq ID No 452; 12-18, 20-25, 43-54, 56-65, 73-79, 82-88, 99-111, 136-142, 153-169, 171-183, 195-223, 229-248, 255-260, 272-277, 281-292, 298-319, 322-329, 332-351, 363-379, 381-389 and 275-304 of Seq ID No 453; 4-9, 34-48, 65-77, 101-106, 111-131, 138-153, 186-191, 230-250 and 148-219 of Seq ID No 454; 4-23, 30-35, 42-53, 67-76, 82-87, 101-108, 112-130, 132-138, 147-152, 161-183, 187-208, 218-225, 265-283, 295-303, 306-317, 322-334, 338-357, 360-368, 370-383, 387-398, 400-419, 421-430, 104-182 and 240-304 of Seq ID No 455; 4-12, 63-69, 94-102, 146-164, 166-173, 175-181, 193-207, 263-281, 286-295, 301-306, 330-343, 369-378, 382-388, 414-420, 422-430, 438-454, 456-462, 472-531, 543-560, 581-591, 596-605, 614-623, 626-635, 656-662, 669-676, 683-690, 693-698, 705-711, 728-736, 752-764 and 69-102 of Seq ID No 456; 6-12, 43-53, 141-147, 164-179, 185-195, 197-206, 227-235, 237-271, 288-305, 308-317, 335-341, 351-357, 365-376, 386-395, 397-416, 422-447 and 11-35 of Seq ID No 457; 16-24, 50-65, 73-84, 88-99, 114-124, 130-146, 181-187, 193-203, 214-220, 236-247, 250-258, 287-297 and 50-113 of Seq ID No 458; 4-25, 50-55, 76-82, 117-123, 131-137, 139-148, 157-166, 239-245, 253-258, 266-275, 277-292, 300-306, 51-83 and 93-161 of Seq ID No 459; 6-22, 34-43, 51-86, 93-100, 110-116, 150-161, 164-171, 180-187, 197-218 and 168-237 of Seq ID No 460; 4-27, 55-60, 74-82 and 10-46 of Seq ID No 461; 6-19, 25-31, 43-49, 60-79, 88-100, 105-129, 148-161, 164-171, 187-193, 243-263, 316-322, 334-340, 369-378, 381-391, 398-404, 460-466, 474-483, 502-509, 511-517, 525-530, 558-567, 579-596, 622-627, 631-638, 641-651, 653-659, 674-680, 687-693, 710-716, 720-727, 743-753, 760-775, 788-795, 806-813, 821-828, 836-842, 847-860, 865-880 and 258-377 of Seq ID No 462; 4-11, 25-64, 71-79, 88-94, 107-120, 123-132, 167-188, 231-237, 240-246, 261-267, 306-311, 330-342, 351-358, 389-395, 406-418, 429-434, 439-448, 483-501, 511-520 and 71-143 of Seq ID No 463; 4-18, 22-27, 53-64, 94-100, 121-127, 133-139, 155-164, 177-

182, 187-196, 206-218, 224-242, 248-253, 258-277 and 184-253 of Seq ID No 464; 10-17, 56-67, 72-82, 94-99, 106-113, 166-173, 229-235, 243-283, 295-301, 313-321, 326-331, 342-348, 396-414, 423-435, 446-452, 454-462, 496-502, 511-534, 543-556, 563-570, 586-593, 616-626, 638-645, 653-662, 679-696, 731-737, 766-774, 776-782, 790-796, 810-817, 825-835, 837-846 and 540-615 of Seq ID No 465; 13-24, 30-36, 73-81, 89-95, 109-115, 117-143, 161-173, 179-189, 226-244, 251-261, 275-281, 298-305, 307-315, 323-328, 364-374, 69-186 and 264-354 of Seq ID No 466; 19-25 and 6-22 of Seq ID No 467; 6-39, 59-68 and 43-62 of Seq ID No 468; 6-14, 22-32 and 1-27 of Seq ID No 469; 4-41 and 28-40 of Seq ID No 470; 8-14 and 4-19 of Seq ID No 471; 4-10, 12-22, 30-35 and 6-33 of Seq ID No 472; 4-16, 24-33 and 37-54 of Seq ID No 473; 2-23 of Seq ID No 474; 4-21, 27-33, 36-41 and 14-34 of Seq ID No 475; 4-14, 24-30, 37-42, 57-78, 83-89, 94-103, 113-131 and 100-122 of Seq ID No 476.

The present invention also provides a process for producing a E. faecalis hyperimmune serum reactive antigen or a fragment thereof according to the present invention comprising expressing one or more of the nucleic acid molecules according to the present invention in a suitable expression system.

Moreover, the present invention provides a process for producing a cell, which expresses a E. faecalis hyperimmune serum reactive antigen or a fragment thereof according to the present invention comprising transforming or transfecting a suitable host cell with the vector according to the present invention.

According to the present invention a pharmaceutical composition, especially a vaccine, comprising a hyperimmune serum-reactive antigen or a fragment thereof as defined in the present invention or a nucleic acid molecule as defined in the present invention is provided.

In a preferred embodiment the pharmaceutical composition further comprises an immunostimulatory substance, preferably selected from the group comprising polycationic polymers, especially polycationic peptides, immunostimulatory deoxynucleotides (ODNs), peptides containing at least two LysLeuLys motifs, especially KLKLLKLK, neuroactive compounds, especially human growth hormone, alumn, Freund's complete or incomplete adjuvants or combinations thereof.

In a more preferred embodiment the immunostimulatory substance is a combination of either a polycationic polymer and immunostimulatory deoxynucleotides or of a peptide containing at least two LysLeuLys motifs and immunostimulatory deoxynucleotides.

In a still more preferred embodiment the polycationic polymer is a polycationic peptide, especially polyarginine.

According to the present invention the use of a nucleic acid molecule according to the present invention or a hyperimmune serum-reactive antigen or fragment thereof according to the present invention for the manufacture of a pharmaceutical preparation, especially for the manufacture of a vaccine against enterococcal infection, is provided.

Also an antibody, or at least an effective part thereof, which binds at least to a selective part of the hyperimmune serum-reactive antigen or a fragment thereof according to the present invention is provided herewith.

In a preferred embodiment the antibody is a monoclonal antibody.

In another preferred embodiment the effective part of the antibody comprises Fab fragments.

In a further preferred embodiment the antibody is a chimeric antibody.

In a still preferred embodiment the antibody is a humanized antibody.

The present invention also provides a hybridoma cell line, which produces an antibody according to the present invention.

Moreover, the present invention provids a method for producing an antibody according to the present invention, characterized by the following steps:
- initiating an immune response in a non-human animal by administrating an hyperimmune serum-reactive antigen or a fragment thereof, as defined in the invention, to said animal,
- removing an antibody containing body fluid from said animal, and
- producing the antibody by subjecting said antibody containing body fluid to further purification steps.

Accordingly, the present invention also provides a method for producing an antibody according to the present invention, characterized by the following steps:
- initiating an immune response in a non-human animal by administrating an hyperimmune serum-reactive antigen or a fragment thereof, as defined in the present invention, to said animal,
- removing the spleen or spleen cells from said animal,
- producing hybridoma cells of said spleen or spleen cells,
- selecting and cloning hybridoma cells specific for said hyperimmune serum-reactive antigens or a fragment thereof,
- producing the antibody by cultivation of said cloned hybridoma cells and optionally further purification steps.

The antibodies provided or produced according to the above methods may be used for the preparation of a medicament for treating or preventing enterococcal infections.

According to another aspect the present invention provides an antagonist which binds to a hyperimmune serum-reactive antigen or a fragment thereof according to the present invention.

Such an antagonist capable of binding to a hyperimmune serum-reactive antigen or fragment thereof according to the present invention may be identified by a method comprising the following steps:
a) contacting an isolated or immobilized hyperimmune serum-reactive antigen or a fragment thereof according to the present invention with a candidate antagonist under conditions to permit binding of said candidate antagonist to said hyperimmune serum-reactive antigen or fragment, in the presence of a component capable of providing a detectable signal in response to the binding of the candidate antagonist to said hyperimmune serum reactive antigen or fragment thereof; and
b) detecting the presence or absence of a signal generated in response to the binding of the antagonist to the hyperimmune serum reactive antigen or the fragment thereof.

An antagonist capable of reducing or inhibiting the interaction activity of a hyperimmune serum-reactive antigen or a fragment thereof according to the present invention to its interaction partner may be identified by a method comprising the following steps:
a) providing a hyperimmune serum reactive antigen or a hyperimmune fragment thereof according to the present invention,
b) providing an interaction partner to said hyperimmune serum reactive antigen or a fragment thereof, especially an antibody according to the present invention,
c) allowing interaction of said hyperimmune serum reactive antigen or fragment thereof to said interaction partner to form an interaction complex,
d) providing a candidate antagonist,

e) allowing a competition reaction to occur between the candidate antagonist and the interaction complex,

f) determining whether the candidate antagonist inhibits or reduces the interaction activities of the hyperimmune serum reactive antigen or the fragment thereof with the interaction partner.

The hyperimmune serum reactive antigens or fragments thereof according to the present invention may be used for the isolation and/or purification and/or identification of an interaction partner of said hyperimmune serum reactive antigen or fragment thereof.

The present invention also provides a process for *in vitro* diagnosing a disease related to expression of a hyperimmune serum-reactive antigen or a fragment thereof according to the present invention comprising determining the presence of a nucleic acid sequence encoding said hyperimmune serum reactive antigen or fragment thereof according to the present invention or the presence of the hyperimmune serum reactive antigen or fragment thereof according to the present invention.

The present invention also provides a process for *in vitro* diagnosis of a bacterial infection, especially a enterococcal infection, comprising analyzing for the presence of a nucleic acid sequence encoding said hyperimmune serum reactive antigen or fragment thereof according to the present invention or the presence of the hyperimmune serum reactive antigen or fragment thereof according to the present invention.

Moreover, the present invention provides the use of a hyperimmune serum reactive antigen or fragment thereof according to the present invention for the generation of a peptide binding to said hyperimmune serum reactive antigen or fragment thereof, wherein the peptide is an anticaline.

The present invention also provides the use of a hyperimmune serum-reactive antigen or fragment thereof according to the present invention for the manufacture of a functional nucleic acid, wherein the functional nucleic acid is selected from the group comprising aptamers and spiegelmers.

The nucleic acid molecule according to the present invention may also be used for the manufacture of a functional ribonucleic acid, wherein the functional ribonucleic acid is selected from the group comprising ribozymes, antisense nucleic acids and siRNA.

The present invention advantageously provides an efficient, relevant and comprehensive set of isolated nucleic acid molecules and their encoded hyperimmune serum reactive antigens or fragments thereof identified from E. faecalis using an antibody preparation from multiple human plasma pools and surface expression libraries derived from the genome of E. faecalis. Thus, the present invention fulfils a widely felt demand for E. faecalis antigens, vaccines, diagnostics and products useful in procedures for preparing antibodies and for identifying compounds effective against enterococcal infections.

An effective vaccine should be composed of proteins or polypeptides, which are expressed by all strains and are able to induce high affinity, abundant antibodies against cell surface components of E. faecalis. The antibodies should be IgG1 and/or IgG3 for opsonization, and any IgG subtype and IgA for neutralisation of adherence and toxin action. A chemically defined vaccine must be definitely superior compared to a whole cell vaccine (attenuated or killed), since components of E. faecalis, which cross-react with human tissues or inhibit opsonization {Whitnack, E. et al., 1985} can be eliminated, and the individual proteins inducing protective antibodies and/or a protective immune response can be selected.

The approach, which has been employed for the present invention, is based on the interaction of enterococcal proteins or peptides with the antibodies present in human sera. The antibodies produced against E. faecalis by the human immune system and present in human sera are indicative of the *in vivo* expression of the antigenic proteins and their immunogenicity. In addition, the antigenic proteins as

identified by the bacterial surface display expression libraries using pools of pre-selected sera, are processed in a second and third round of screening by individual selected or generated sera. Thus the present invention supplies an efficient, relevant, comprehensive set of enterococcal antigens as a pharmaceutical composition, especially a vaccine preventing infection by E. faecalis.

In the antigen identification program for identifying a comprehensive set of antigens according to the present invention, at least two different bacterial surface expression libraries are screened with several serum pools or plasma fractions or other pooled antibody containing body fluids (antibody pools). The antibody pools are derived from a serum collection, which has been tested against antigenic compounds of E. faecalis, such as whole cell extracts and culture supernatant proteins. Preferably, 2 distinct serum collections are used: 1. With very stable antibody repertoire: normal adults, clinically healthy people, who are non-carriers and overcame previous encounters or currently carriers of E. faecalis without acute disease and symptoms, 2. With antibodies induced acutely by the presence of the pathogenic organism: patients with acute disease with different manifestations (e.g. E. faecalis endocarditis, urinary tract infection and bacteraemia). Sera have to react with multiple enterococci-specific antigens in order to be considered hyperimmune and therefore relevant in the screening method applied for the present invention. The antibodies produced against enterococci by the human immune system and present in human sera are indicative of the in vivo expression of the antigenic proteins and their immunogenicity.

The expression libraries as used in the present invention should allow expression of all potential antigens, e.g. derived from all surface proteins of E. faecalis. Bacterial surface display libraries will be represented by a recombinant library of a bacterial host displaying a (total) set of expressed peptide sequences of enterococci on a number of selected outer membrane proteins (LamB, BtuB, FhuA) at the bacterial host membrane (Georgiou, G., 1997); (Etz, H. et al., 2001). One of the advantages of using recombinant expression libraries is that the identified hyperimmune serum-reactive antigens may be instantly produced by expression of the coding sequences of the screened and selected clones expressing the hyperimmune serum-reactive antigens without further recombinant DNA technology or cloning steps necessary.

The comprehensive set of antigens identified by the described program according to the present invention is analysed further by one or more additional rounds of screening. Therefore individual antibody preparations or antibodies generated against selected peptides which were identified as immunogenic are used. According to a preferred embodiment the individual antibody preparations for the second round of screening are derived from patients who have suffered from an acute infection with enterococci, especially from patients who show an antibody titer above a certain minimum level, for example an antibody titer being higher than 80 percentile, preferably higher than 90 percentile, especially higher than 95 percentile of the human (patient or healthy individual) sera tested. Using such high titer individual antibody preparations in the second screening round allows a very selective identification of the hyperimmune serum-reactive antigens and fragments thereof from E. faecalis.

Following the comprehensive screening procedure, the selected antigenic proteins, expressed as recombinant proteins or in vitro translated products, in case it can not be expressed in prokaryotic expression systems, or the identified antigenic peptides (produced synthetically) are tested in a second screening by a series of ELISA and Western blotting assays for the assessment of their immunogenicity with a large human serum collection (> 100 uninfected, > 50 patients sera).

It is important that the individual antibody preparations (which may also be the selected serum) allow a selective identification of the most promising candidates of all the hyperimmune serum-reactive antigens from all the promising candidates from the first round. Therefore, preferably at least 10 individual antibody preparations (i.e. antibody preparations (e.g. sera) from at least 10 different individuals having suffered from an infection to the chosen pathogen) should be used in identifying these antigens in the second screening round. Of course, it is possible to use also less than 10 individual preparations,

however, selectivity of the step may not be optimal with a low number of individual antibody preparations. On the other hand, if a given hyperimmune serum-reactive antigen (or an antigenic fragment thereof) is recognized by at least 10 individual antibody preparations, preferably at least 30, especially at least 50 individual antibody preparations, identification of the hyperimmune serum-reactive antigen is also selective enough for a proper identification. Hyperimmune serum-reactivity may of course be tested with as many individual preparations as possible (e.g. with more than 100 or even with more than 1,000).

Therefore, the relevant portion of the hyperimmune serum-reactive antibody preparations according to the method of the present invention should preferably be at least 10, more preferred at least 30, especially at least 50 individual antibody preparations. Alternatively (or in combination) hyperimmune serum-reactive antigens may preferably be also identified with at least 20%, preferably at least 30%, especially at least 40% of all individual antibody preparations used in the second screening round.

According to a preferred embodiment of the present invention, the sera from which the individual antibody preparations for the second round of screening are prepared (or which are used as antibody preparations), are selected by their titer against *E. faecalis* (e.g. against a preparation of this pathogen, such as a lysate, cell wall components and recombinant proteins). Preferably, some are selected with a total IgG titer above 10,000 U, especially above 12,000 U (U = units, calculated from the OD$_{405nm}$ reading at a given dilution) when the whole organism (total lysate or whole cells) is used as antigen in the ELISA.

The antibodies produced against enterococci by the human immune system and present in human sera are indicative of the in vivo expression of the antigenic proteins and their immunogenicity. The recognition of linear epitopes by antibodies can be based on sequences as short as 4-5 amino acids. Of course it does not necessarily mean that these short peptides are capable of inducing the given antibody in vivo. For that reason the defined epitopes, polypeptides and proteins are further to be tested in animals (mainly in mice) for their capacity to induce antibodies against the selected proteins in vivo.

The preferred antigens are located on the cell surface or are secreted, and are therefore accessible extracellularly. Antibodies against cell wall proteins are expected to serve two purposes: to inhibit adhesion and to promote phagocytosis. Antibodies against secreted proteins are beneficial in neutralisation of their function as toxin or virulence component. It is also known that bacteria communicate with each other through secreted proteins. Neutralizing antibodies against these proteins will interrupt growth-promoting cross-talk between or within enterococcal species. Bioinformatic analyses (signal sequences, cell wall localisation signals, transmembrane domains) proved to be very useful in assessing cell surface localisation or secretion. The experimental approach includes the isolation of antibodies with the corresponding epitopes from human serum, and the generation of immune sera in mice against (poly)peptides selected by the bacterial surface display screens.

For that purpose, bacterial *E. coli* clones are directly injected into mice and immune sera are taken and tested in the relevant in vitro assay for functional opsonic or neutralizing antibodies. Alternatively, specific antibodies may be purified from human or mouse sera using peptides or proteins as substrate.

Host defence against *E. faecalis* relies mainly on innate immunological mechanisms. Inducing high affinity antibodies of the opsonic and neutralizing type by vaccination helps the innate immune system to eliminate bacteria and toxins. This makes the method according to the present invention an optimal tool for the identification of enterococcal antigenic proteins.

The skin and mucous membranes are formidable barriers against invasion by enterococci. However, once the skin or the mucous membranes are breached the first line of non-adaptive cellular defence begins its co-ordinate action through complement and phagocytes, especially the polymorphonuclear leukocytes (PMNs). These cells can be regarded as the cornerstones in eliminating invading bacteria. As enterococci

are primarily extracellular pathogens, the major anti-enterococcal adaptive response comes from the humoral arm of the immune system, and is mediated through three major mechanisms: promotion of opsonization, toxin neutralisation, and inhibition of adherence. It is believed that opsonization is especially important, because of its requirement for an effective phagocytosis. For efficient opsonization the microbial surface has to be coated with antibodies and complement factors for recognition by PMNs through receptors to the Fc fragment of the IgG molecule or to activated C3b. After opsonization, enterococci are phagocytosed and killed. Antibodies bound to specific antigens on the cell surface of bacteria serve as ligands for the attachment to PMNs and to promote phagocytosis. The very same antibodies bound to the adhesins and other cell surface proteins are expected to neutralize adhesion and prevent colonization. The selection of antigens as provided by the present invention is thus well suited to identify those that will lead to protection against infection in an animal model or in humans.

According to the antigen identification method used herein, the present invention can surprisingly provide a set of comprehensive novel nucleic acids and novel hyperimmune serum reactive antigens and fragments thereof of *E. faecalis*, among other things, as described below. According to one aspect, the invention particularly relates to the nucleotide sequences encoding hyperimmune serum reactive antigens which sequences are set forth in the Sequence listing Seq ID No 1-170, 373-424 and the corresponding encoded amino acid sequences representing hyperimmune serum reactive antigens are set forth in the Sequence Listing Seq ID No 171-340 and 425-476.

In a preferred embodiment of the present invention, a nucleic acid molecule is provided which exhibits 70% identity over their entire length to a nucleotide sequence set forth with Seq ID No 1-2, 4-8, 10, 12-18, 20-23, 25-26, 29-43, 45-62, 64-74, 76-77, 79-83, 85-89, 91-92, 94-114, 117-126, 128-146, 148-170, 373, 375, 379-381, 387, 392, 394, 397-399, 407-408, 410-411 and 415-424. Most highly preferred are nucleic acids that comprise a region that is at least 80% or at least 85% identical over their entire length to a nucleic acid molecule set forth with Seq ID No 1-2, 4-8, 10, 12-18, 20-23, 25-26, 29-43, 45-62, 64-74, 76-77, 79-83, 85-89, 91-92, 94-114, 117-126, 128-146, 148-170, 373, 375, 379-381, 387, 392, 394, 397-399, 407-408, 410-411 and 415-424. In this regard, nucleic acid molecules at least 90%, 91%, 92%, 93%, 94%, 95%, or 96% identical over their entire length to the same are particularly preferred. Furthermore, those with at least 97% are highly preferred, those with at least 98% and at least 99% are particularly highly preferred, with at least 99% or 99.5% being the more preferred, with 100% identity being especially preferred. Moreover, preferred embodiments in this respect are nucleic acids which encode hyperimmune serum reactive antigens or fragments thereof (polypeptides) which retain substantially the same biological function or activity as the mature polypeptide encoded by said nucleic acids set forth in the Seq ID No 1-2, 4-8, 10, 12-18, 20-23, 25-26, 29-43, 45-62, 64-74, 76-77, 79-83, 85-89, 91-92, 94-114, 117-126, 128-146, 148-170, 373, 375, 379-381, 387, 392, 394, 397-399, 407-408, 410-411 and 415-424.

Identity, as known in the art and used herein, is the relationship between two or more polypeptide sequences or two or more polynucleotide sequences, as determined by comparing the sequences. In the art, identity also means the degree of sequence relatedness between polypeptide or polynucleotide sequences, as the case may be, as determined by the match between strings of such sequences. Identity can be readily calculated. While there exist a number of methods to measure identity between two polynucleotide or two polypeptide sequences, the term is well known to skilled artisans (e.g. *Sequence Analysis in Molecular Biology*, von Heinje, G., Academic Press, 1987). Preferred methods to determine identity are designed to give the largest match between the sequences tested. Methods to determine identity are codified in computer programs. Preferred computer program methods to determine identity between two sequences include, but are not limited to, GCG program package (Devereux, J. et al., 1984), BLASTP, BLASTN, and FASTA (Altschul, S. et al., 1990).

According to another aspect of the invention, nucleic acid molecules are provided which exhibit at least 96%, preferably at least 98 %, especially 100 % identity to the nucleic acid sequence set forth with Seq ID No 3, 9, 11, 24, 27, 44, 63, 75, 84, 115-116, 127, 374, 376-378, 382-386, 388-391, 393, 395-396, 400-406, 409 and

412-414.

According to a further aspect of the present invention, nucleic acid molecules are provided which are identical to the nucleic acid sequences set forth with Seq ID No 90, 147.

The nucleic acid molecules according to the present invention can as a second alternative also be a nucleic acid molecule which is at least essentially complementary to the nucleic acid described as the first alternative above. As used herein complementary means that a nucleic acid strand is base pairing via Watson-Crick base pairing with a second nucleic acid strand. Essentially complementary as used herein means that the base pairing is not occurring for all of the bases of the respective strands but leaves a certain number or percentage of the bases unpaired or wrongly paired. The percentage of correctly pairing bases is preferably at least 70 %, more preferably 80 %, even more preferably 90 % and most preferably any percentage higher than 90 %. It is to be noted that a percentage of 70 % matching bases is considered as homology and the hybridization having this extent of matching base pairs is considered as stringent. Hybridization conditions for this kind of stringent hybridization may be taken from Current Protocols in Molecular Biology (John Wiley and Sons, Inc., 1987). More particularly, the hybridization conditions can be as follows:

- Hybridization performed e.g. in 5 x SSPE, 5 x Denhardt's reagent, 0.1% SDS, 100 g/mL sheared DNA at 68°C
- Moderate stringency wash in 0.2xSSC, 0.1% SDS at 42°C
- High stringency wash in 0.1xSSC, 0.1% SDS at 68°C

Genomic DNA with a GC content of 50% has an approximate T_M of 96°C. For 1% mismatch, the T_M is reduced by approximately 1°C.

In addition, any of the further hybridization conditions described herein are in principle applicable as well.

Of course, all nucleic acid sequence molecules which encode the same polypeptide molecule as those identified by the present invention are encompassed by any disclosure of a given coding sequence, since the degeneracy of the genetic code is directly applicable to unambiguously determine all possible nucleic acid molecules which encode a given polypeptide molecule, even if the number of such degenerated nucleic acid molecules may be high. This is also applicable for fragments of a given polypeptide, as long as the fragments encode a polypeptide being suitable to be used in a vaccination connection, e.g. as an active or passive vaccine.

The nucleic acid molecule according to the present invention can as a third alternative also be a nucleic acid which comprises a stretch of at least 15 bases of the nucleic acid molecule according to the first and second alternative of the nucleic acid molecules according to the present invention as outlined above. Preferably, the bases form a contiguous stretch of bases. However, it is also within the scope of the present invention that the stretch consists of two or more moieties which are separated by a number of bases.

The present nucleic acids may preferably consist of at least 20, even more preferred at least 30, especially at least 50 contiguous bases from the sequences disclosed herein. The suitable length may easily be optimized due to the planned area of use (e.g. as (PCR) primers, probes, capture molecules (e.g. on a (DNA) chip), etc.). Preferred nucleic acid molecules contain at least a contiguous 15 base portion of one or more of the predicted immunogenic amino acid sequences listed in tables 1 and 2, especially the sequences of table 2 with scores of more than 10, preferably more than 20, especially with a score of more than 25. Specifically preferred are nucleic acids containing a contiguous portion of a DNA sequence of any sequence in the sequence protocol of the present application which shows 1 or more, preferably more

than 2, especially more than 5, non-identical nucleic acid residues compared to the published *Enterococcus faecalis* strain V583 genome {Paulsen, I. et al., 2003}; GenBank accession AE016830 (chromosome), AE016833 (pTEF1), AE016831 (pTEF2), AE016832 (pTEF3), and/or any other published *E. faecalis* genome sequence or parts thereof. Specifically preferred non-identical nucleic acid residues are residues which lead to a non-identical amino acid residue. Preferably, the nucleic acid sequences encode for polypeptides having at least 1, preferably at least 2, preferably at least three different amino acid residues compared to the published S.pyogenes counterparts mentioned above. Also such isolated polypeptides, being fragments of the proteins (or the whole protein) mentioned herein e.g. in the sequence listing, having at least 6, 7, or 8 amino acid residues and being encoded by these nucleic acids are preferred.

The nucleic acid molecule according to the present invention can as a fourth alternative also be a nucleic acid molecule which anneals under stringent hybridisation conditions to any of the nucleic acids of the present invention according to the above outlined first, second, and third alternative. Stringent hybridisation conditions are typically those described herein.

Finally, the nucleic acid molecule according to the present invention can as a fifth alternative also be a nucleic acid molecule which, but for the degeneracy of the genetic code, would hybridise to any of the nucleic acid molecules according to any nucleic acid molecule of the present invention according to the first, second, third, and fourth alternative as outlined above. This kind of nucleic acid molecule refers to the fact that preferably the nucleic acids according to the present invention code for the hyperimmune serum reactive antigens or fragments thereof according to the present invention. This kind of nucleic acid molecule is particularly useful in the detection of a nucleic acid molecule according to the present invention and thus the diagnosis of the respective microorganisms such as *E. faecalis* and any disease or diseased condition where this kind of microorganims is involved. Preferably, the hybridisation would occur or be preformed under stringent conditions as described in connection with the fourth alternative described above.

Nucleic acid molecule as used herein generally refers to any ribonucleic acid molecule or deoxyribonucleic acid molecule, which may be unmodified RNA or DNA or modified RNA or DNA. Thus, for instance, nucleic acid molecule as used herein refers to, among other, single-and double-stranded DNA, DNA that is a mixture of single- and double-stranded RNA, and RNA that is a mixture of single- and double-stranded regions, hybrid molecules comprising DNA and RNA that may be single-stranded or, more typically, double-stranded, or triple-stranded, or a mixture of single- and double-stranded regions. In addition, nucleic acid molecule as used herein refers to triple-stranded regions comprising RNA or DNA or both RNA and DNA. The strands in such regions may be from the same molecule or from different molecules. The regions may include all of one or more of the molecules, but more typically involve only a region of some of the molecules. One of the molecules of a triple-helical region often is an oligonucleotide. As used herein, the term nucleic acid molecule includes DNAs or RNAs as described above that contain one or more modified bases. Thus, DNAs or RNAs with backbones modified for stability or for other reasons are "nucleic acid molecule" as that term is intended herein. Moreover, DNAs or RNAs comprising unusual bases, such as inosine, or modified bases, such as tritylated bases, to name just two examples, are nucleic acid molecule as the term is used herein. It will be appreciated that a great variety of modifications have been made to DNA and RNA that serve many useful purposes known to those of skill in the art. The term nucleic acid molecule as it is employed herein embraces such chemically, enzymatically or metabolically modified forms of nucleic acid molecule, as well as the chemical forms of DNA and RNA characteristic of viruses and cells, including simple and complex cells, *inter alia*. The term nucleic acid molecule also embraces short nucleic acid molecules often referred to as oligonucleotide(s). "Polynucleotide" and "nucleic acid" or "nucleic acid molecule" are often used interchangeably herein.

Nucleic acid molecules provided in the present invention also encompass numerous unique fragments, both longer and shorter than the nucleic acid molecule sequences set forth in the sequencing listing of the

E. faecalis coding regions, which can be generated by standard cloning methods. To be unique, a fragment must be of sufficient size to distinguish it from other known nucleic acid sequences, most readily determined by comparing any selected *E. faecalis* fragment to the nucleotide sequences in computer databases such as GenBank.

Additionally, modifications can be made to the nucleic acid molecules and polypeptides that are encompassed by the present invention. For example, nucleotide substitutions can be made which do not affect the polypeptide encoded by the nucleic acid, and thus any nucleic acid molecule which encodes a hyperimmune serum reactive antigen or fragments thereof is encompassed by the present invention.

Furthermore, any of the nucleic acid molecules encoding hyperimmune serum reactive antigens or fragments thereof provided by the present invention can be functionally linked, using standard techniques such as standard cloning techniques, to any desired regulatory sequences, whether a *E. faecalis* regulatory sequence or a heterologous regulatory sequence, heterologous leader sequence, heterologous marker sequence or a heterologous coding sequence to create a fusion protein.

Nucleic acid molecules of the present invention may be in the form of RNA, such as mRNA or cRNA, or in the form of DNA, including, for instance, cDNA and genomic DNA obtained by cloning or produced by chemical synthetic techniques or by a combination thereof. The DNA may be triple-stranded, double-stranded or single-stranded. Single-stranded DNA may be the coding strand, also known as the sense strand, or it may be the non-coding strand, also referred to as the anti-sense strand.

The present invention further relates to variants of the herein above described nucleic acid molecules which encode fragments, analogs and derivatives of the hyperimmune serum reactive antigens and fragments thereof having a deducted *E. faecalis* amino acid sequence set forth in the Sequence Listing. A variant of the nucleic acid molecule may be a naturally occurring variant such as a naturally occurring allelic variant, or it may be a variant that is not known to occur naturally. Such non-naturally occurring variants of the nucleic acid molecule may be made by mutagenesis techniques, including those applied to nucleic acid molecules, cells or organisms.

Among variants in this regard are variants that differ from the aforementioned nucleic acid molecules by nucleotide substitutions, deletions or additions. The substitutions, deletions or additions may involve one or more nucleotides. The variants may be altered in coding or non-coding regions or both. Alterations in the coding regions may produce conservative or non-conservative amino acid substitutions, deletions or additions. Preferred are nucleic acid molecules encoding a variant, analog, derivative or fragment, or a variant, analogue or derivative of a fragment, which have a *E. faecalis* sequence as set forth in the Sequence Listing, in which several, a few, 5 to 10, 1 to 5, 1 to 3, 2, 1 or no amino acid(s) is substituted, deleted or added, in any combination. Especially preferred among these are silent substitutions, additions and deletions, which do not alter the properties and activities of the *E. faecalis* polypeptides set forth in the Sequence Listing. Also especially preferred in this regard are conservative substitutions.

The peptides and fragments according to the present invention also include modified epitopes wherein preferably one or two of the amino acids of a given epitope are modified or replaced according to the rules disclosed in e.g. {Tourdot, S. et al., 2000}, as well as the nucleic acid sequences encoding such modified epitopes.

It is clear that also epitopes derived from the present epitopes by amino acid exchanges improving, conserving or at least not significantly impeding the T cell activating capability of the epitopes are covered by the epitopes according to the present invention. Therefore the present epitopes also cover epitopes, which do not contain the original sequence as derived from *E. faecalis*, but trigger the same or preferably an improved T cell response. These epitope are referred to as "heteroclitic"; they need to have a

similar or preferably greater affinity to MHC/HLA molecules, and the need the ability to stimulate the T cell receptors (TCR) directed to the original epitope in a similar or preferably stronger manner.

Heteroclitic epitopes can be obtained by rational design i.e. taking into account the contribution of individual residues to binding to MHC/HLA as for instance described by {Rammensee, H. et al., 1999}, combined with a systematic exchange of residues potentially interacting with the TCR and testing the resulting sequences with T cells directed against the original epitope. Such a design is possible for a skilled man in the art without much experimentation.

Another possibility includes the screening of peptide libraries with T cells directed against the original epitope. A preferred way is the positional scanning of synthetic peptide libraries. Such approaches have been described in detail for instance by {Hemmer, B. et al., 1999} and the references given therein.

As an alternative to epitopes represented by the present derived amino acid sequences or heteroclitic epitopes, also substances mimicking these epitopes e.g. "peptidemimetica" or "retro-inverso-peptides" can be applied.

Another aspect of the design of improved epitopes is their formulation or modification with substances increasing their capacity to stimulate T cells. These include T helper cell epitopes, lipids or liposomes or preferred modifications as described in WO 01/78767.

Another way to increase the T cell stimulating capacity of epitopes is their formulation with immune stimulating substances for instance cytokines or chemokines like interleukin-2, -7, -12, -18, class I and II interferons (IFN), especially IFN-gamma, GM-CSF, TNF-alpha, flt3-ligand and others.

As discussed additionally herein regarding nucleic acid molecule assays of the invention, for instance, nucleic acid molecules of the invention as discussed above, may be used as a hybridization probe for RNA, cDNA and genomic DNA to isolate full-length cDNAs and genomic clones encoding polypeptides of the present invention and to isolate cDNA and genomic clones of other genes that have a high sequence similarity to the nucleic acid molecules of the present invention. Such probes generally will comprise at least 15 bases. Preferably, such probes will have at least 20, at least 25 or at least 30 bases, and may have at least 50 bases. Particularly preferred probes will have at least 30 bases, and will have 50 bases or less, such as 30, 35, 40, 45, or 50 bases.

For example, the coding region of a nucleic acid molecule of the present invention may be isolated by screening a relevant library using the known DNA sequence to synthesize an oligonucleotide probe. A labeled oligonucleotide having a sequence complementary to that of a gene of the present invention is then used to screen a library of cDNA, genomic DNA or mRNA to determine to which members of the library the probe hybridizes.

The nucleic acid molecules and polypeptides of the present invention may be employed as reagents and materials for development of treatments of and diagnostics for disease, particularly human disease, as further discussed herein relating to nucleic acid molecule assays, inter alia.

The nucleic acid molecules of the present invention that are oligonucleotides can be used in the processes herein as described, but preferably for PCR, to determine whether or not the E. faecalis genes identified herein in whole or in part are present and/or transcribed in infected tissue such as blood. It is recognized that such sequences will also have utility in diagnosis of the stage of infection and type of infection the pathogen has attained. For this and other purposes the arrays comprising at least one of the nucleic acids according to the present invention as described herein, may be used.

- 23 -

The nucleic acid molecules according to the present invention may be used for the detection of nucleic acid molecules and organisms or samples containing these nucleic acids. Preferably such detection is for diagnosis, more preferable for the diagnosis of a disease related or linked to the present or abundance of *E. faecalis*.

Eukaryotes (herein also "individual(s)"), particularly mammals, and especially humans, infected with *E. faecalis* may be identifiable by detecting any of the nucleic acid molecules according to the present invention detected at the DNA level by a variety of techniques. Preferred nucleic acid molecules candidates for distinguishing a *E. faecalis* from other organisms can be obtained.

The invention provides a process for diagnosing disease, arising from infection with *E. faecalis*, comprising determining from a sample isolated or derived from an individual an increased level of expression of a nucleic acid molecule having the sequence of a nucleic acid molecule set forth in the Sequence Listing. Expression of nucleic acid molecules can be measured using any one of the methods well known in the art for the quantitation of nucleic acid molecules, such as, for example, PCR, RT-PCR, Rnase protection, Northern blotting, other hybridisation methods and the arrays described herein.

Isolated as used herein means separated "by the hand of man" from its natural state; i.e., that, if it occurs in nature, it has been changed or removed from its original environment, or both. For example, a naturally occurring nucleic acid molecule or a polypeptide naturally present in a living organism in its natural state is not "isolated," but the same nucleic acid molecule or polypeptide separated from the coexisting materials of its natural state is "isolated", as the term is employed herein. As part of or following isolation, such nucleic acid molecules can be joined to other nucleic acid molecules, such as DNAs, for mutagenesis, to form fusion proteins, and for propagation or expression in a host, for instance. The isolated nucleic acid molecules, alone or joined to other nucleic acid molecules such as vectors, can be introduced into host cells, in culture or in whole organisms. Introduced into host cells in culture or in whole organisms, such DNAs still would be isolated, as the term is used herein, because they would not be in their naturally occurring form or environment. Similarly, the nucleic acid molecules and polypeptides may occur in a composition, such as a media formulations, solutions for introduction of nucleic acid molecules or polypeptides, for example, into cells, compositions or solutions for chemical or enzymatic reactions, for instance, which are not naturally occurring compositions, and, therein remain isolated nucleic acid molecules or polypeptides within the meaning of that term as it is employed herein.

The nucleic acids according to the present invention may be chemically synthesized. Alternatively, the nucleic acids can be isolated from *E. faecalis* by methods known to the one skilled in the art.

According to another aspect of the present invention, a comprehensive set of novel hyperimmune serum reactive antigens and fragments thereof are provided by using the herein described antigen identification method. In a preferred embodiment of the invention, a hyperimmune serum-reactive antigen comprising an amino acid sequence being encoded by any one of the nucleic acids molecules herein described and fragments thereof are provided. In another preferred embodiment of the invention a novel set of hyperimmune serum-reactive antigens which comprises amino acid sequences selected from a group consisting of the polypeptide sequences as represented in Seq ID No 171-172, 174-178, 180, 182-188, 190-193, 195-196, 199-213, 215-232, 234-244, 246-247, 249-253, 255-259, 261-262, 264-284, 287-296, 298-316, 318-340, 425, 427, 431-433, 439, 444, 446, 449-451, 459-460, 462-463 and 467-476 and fragments thereof are provided. In a further preferred embodiment of the invention hyperimmune serum-reactive antigens which comprise amino acid sequences selected from a group consisting of the polypeptide sequences as represented in Seq ID No 173, 179, 181, 194, 197, 214, 233, 245, 254, 285-286, 297, 426, 428-430, 434-438, 440-443, 445, 447-448, 452-458, 461 and 464-466 and fragments thereof are provided. In a still preferred embodiment of the invention hyperimmune serum-reactive antigens which comprise amino acid sequences selected from a group consisting of the polypeptide sequences as represented in Seq ID No 260, 317 and fragments thereof are provided.

The hyperimmune serum reactive antigens and fragments thereof as provided in the invention include any polypeptide set forth in the Sequence Listing as well as polypeptides which have at least 70% identity to a polypeptide set forth in the Sequence Listing, preferably at least 80% or 85% identity to a polypeptide set forth in the Sequence Listing, and more preferably at least 90% similarity (more preferably at least 90% identity) to a polypeptide set forth in the Sequence Listing and still more preferably at least 95%, 96%, 97%, 98%, 99% or 99.5% similarity (still more preferably at least 95%, 96%, 97%, 98%, 99%, or 99.5% identity) to a polypeptide set forth in the Sequence Listing and also include portions of such polypeptides with such portion of the polypeptide generally containing at least 4 amino acids and more preferably at least 8, still more preferably at least 30, still more preferably at least 50 amino acids, such as 4, 8, 10, 20, 30, 35, 40, 45 or 50 amino acids.

The invention also relates to fragments, analogs, and derivatives of these hyperimmune serum reactive antigens and fragments thereof. The terms "fragment", "derivative" and "analog" when referring to an antigen whose amino acid sequence is set forth in the Sequence Listing, means a polypeptide which retains essentially the same or a similar biological function or activity as such hyperimmune serum reactive antigen and fragment thereof.

The fragment, derivative or analog of a hyperimmune serum reactive antigen and fragment thereof may be 1) one in which one or more of the amino acid residues are substituted with a conserved or non-conserved amino acid residue (preferably a conserved amino acid residue) and such substituted amino acid residue may or may not be one encoded by the genetic code, or 2) one in which one or more of the amino acid residues includes a substituent group, or 3) one in which the mature hyperimmune serum reactive antigen or fragment thereof is fused with another compound, such as a compound to increase the half-life of the hyperimmune serum reactive antigen and fragment thereof (for example, polyethylene glycol), or 4) one in which the additional amino acids are fused to the mature hyperimmune serum reactive antigen or fragment thereof, such as a leader or secretory sequence or a sequence which is employed for purification of the mature hyperimmune serum reactive antigen or fragment thereof or a proprotein sequence. Such fragments, derivatives and analogs are deemed to be within the scope of those skilled in the art from the teachings herein.

The present invention also relates to antigens of different E. faecalis isolates. Such homologues may easily be isolated based on the nucleic acid and amino acid sequences disclosed herein. The presence of any antigen can accordingly be determined for every M serotype. In addition it is possible to determine the variability of a particular antigen in the various serotypes as described for the sic gene (Hoe, N. et al., 2001). The contriubtion of the various serotypes to the different enterococcal infections varies in the different age groups and geographical regions. It is an important aspect that the most valuable protective antigens are expected to be conserved among various clinical strains.

Among the particularly preferred embodiments of the invention in this regard are the hyperimmune serum reactive antigens set forth in the Sequence Listing, variants, analogs, derivatives and fragments thereof, and variants, analogs and derivatives of fragments. Additionally, fusion polypeptides comprising such hyperimmune serum reactive antigens, variants, analogs, derivatives and fragments thereof, and variants, analogs and derivatives of the fragments are also encompassed by the present invention. Such fusion polypeptides and proteins, as well as nucleic acid molecules encoding them, can readily be made using standard techniques, including standard recombinant techniques for producing and expression of a recombinant polynucleic acid encoding a fusion protein.

Among preferred variants are those that vary from a reference by conservative amino acid substitutions. Such substitutions are those that substitute a given amino acid in a polypeptide by another amino acid of like characteristics. Typically seen as conservative substitutions are the replacements, one for another,

among the aliphatic amino acids Ala, Val, Leu and Ile; interchange of the hydroxyl residues Ser and Thr, exchange of the acidic residues Asp and Glu, substitution between the amide residues Asn and Gln, exchange of the basic residues Lys and Arg and replacements among the aromatic residues Phe and Tyr.

Further particularly preferred in this regard are variants, analogs, derivatives and fragments, and variants, analogs and derivatives of the fragments, having the amino acid sequence of any polypeptide set forth in the Sequence Listing, in which several, a few, 5 to 10, 1 to 5, 1 to 3, 2, 1 or no amino acid residues are substituted, deleted or added, in any combination. Especially preferred among these are silent substitutions, additions and deletions, which do not alter the properties and activities of the polypeptide of the present invention. Also especially preferred in this regard are conservative substitutions. Most highly preferred are polypeptides having an amino acid sequence set forth in the Sequence Listing without substitutions.

The hyperimmune serum reactive antigens and fragments thereof of the present invention are preferably provided in an isolated form, and preferably are purified to homogeneity.

Also among preferred embodiments of the present invention are polypeptides comprising fragments of the polypeptides having the amino acid sequence set forth in the Sequence Listing, and fragments of variants and derivatives of the polypeptides set forth in the Sequence Listing.

In this regard a fragment is a polypeptide having an amino acid sequence that entirely is the same as part but not all of the amino acid sequence of the afore mentioned hyperimmune serum reactive antigen and fragment thereof, and variants or derivative, analogs, fragments thereof. Such fragments may be "free-standing", i.e., not part of or fused to other amino acids or polypeptides, or they may be comprised within a larger polypeptide of which they form a part or region. Also preferred in this aspect of the invention are fragments characterised by structural or functional attributes of the polypeptide of the present invention, i.e. fragments that comprise alpha-helix and alpha-helix forming regions, beta-sheet and beta-sheet forming regions, turn and turn-forming regions, coil and coil-forming regions, hydrophilic regions, hydrophobic regions, alpha amphipathic regions, beta-amphipathic regions, flexible regions, surface-forming regions, substrate binding regions, and high antigenic index regions of the polypeptide of the present invention, and combinations of such fragments. Preferred regions are those that mediate activities of the hyperimmune serum reactive antigens and fragments thereof of the present invention. Most highly preferred in this regard are fragments that have a chemical, biological or other activity of the hyperimmune serum reactive antigen and fragments thereof of the present invention, including those with a similar activity or an improved activity, or with a decreased undesirable activity. Particularly preferred are fragments comprising receptors or domains of enzymes that confer a function essential for viability of *E. faecalis* or the ability to cause disease in humans. Further preferred polypeptide fragments are those that comprise or contain antigenic or immunogenic determinants in an animal, especially in a human.

An antigenic fragment is defined as a fragment of the identified antigen which is for itself antigenic or may be made antigenic when provided as a hapten. Therefore, also antigens or antigenic fragments showing one or (for longer fragments) only a few amino acid exchanges are enabled with the present invention, provided that the antigenic capacities of such fragments with amino acid exchanges are not severely deteriorated on the exchange(s), i.e., suited for eliciting an appropriate immune response in an individual vaccinated with this antigen and identified by individual antibody preparations from individual sera.

Preferred examples of such fragments of a hyperimmune serum-reactive antigen are selected from the group consisting of peptides comprising amino acid sequences of column "predicted immunogenic aa", and "Location of identified immunogenic region" of Table 1a and Table 1c; the serum reactive epitopes of Table 2, especially peptides comprising amino acid 4-10, 14-21, 30-36, 59-68, 77-82, 87-93, 96-105, 112-121,

125-133, 135-141, 150-162, 164-183, 192-203, 207-213, 215-226, 228-234, 241-247, 250-285, 302-308 and 135-148 of Seq ID No 171; 15-57, 60-73, 77-101, 108-134, 136-177, 185-201, 203-217, 226-240, 244-254, 272-277, 283-288, 292-343, 354-370, 380-398, 406-437, 439-453, 473-490, 532-538, 584-590, 595-601, 606-612, 664-677, 679-704, 715-724, 731-753, 759-772, 786-794, 814-862 and 657-684 of Seq ID No 172; 4-9, 15-36, 41-47, 54-60, 75-81, 114-120, 131-146, 152-158, 174-182, 194-202, 208-215, 218-226, 255-271, 276-285, 290-295, 302-311, 318-328, 330-344, 352-359, 365-377, 388-395, 398-405, 426-432, 439-449, 455-500, 505-513, 531-537, 542-552, 554-561, 587-595, 606-612, 718-734, 763-771, 775-782, 792-801, 805-812, 822-828, 830-843, 849-863, 876-894, 905-911, 919-926, 935-947, 949-958, 968-979, 1009-1016, 1029-1045, 1047-1056, 1076-1081, 1092-1106, 1123-1133, 1179-1200, 1202-1211, 1215-1223, 1287-1299, 1301-1306, 398-431 and 1224-1237 of Seq ID No 173; 17-47, 74-80, 90-97, 126-133, 137-148, 167-173, 179-185, 214-223, 250-255, 270-283, 329-338, 342-350, 352-358, 360-367, 372-383, 398-404, 411-421, 426-432, 435-446, 452-462, 472-479, 515-521, 582-592, 611-618, 623-629, 642-659, 666-673, 678-689, 704-725, 732-737, 744-757, 768-789, 824-834, 842-849, 862-868, 877-887, 904-916, 923-928, 941-947, 962-974, 982-992, 1019-1030, 1032-1044, 1046-1052, 1065-1075, 1077-1087, 1108-1121, 1124-1132, 1137-1151, 1170-1182, 1190-1206, 1208-1214, 1227-1233, 1242-1251, 1254-1273, 1282-1298 and 792-825 of Seq ID No 174; 19-31, 39-67, 82-91, 104-110, 113-128, 149-155, 161-181 and 137-155 of Seq ID No 175; 6-18, 54-63, 69-85, 110-127, 142-156, 158-167, 169-211, 238-246, 248-257, 276-311, 339-349, 371-380, 385-391, 394-403, 421-438, 451-456, 483-489 and 449-468 of Seq ID No 176; 5-15, 24-34, 50-56, 61-83, 98-121, 123-136, 149-162, 166-194, 202-215, 221-227, 229-332, 337-360, 367-402, 404-415, 427-433, 444-462, 471-478, 487-498, 511-518, 521-544, 550-563, 568-574, 580-587, 597-607, 610-616, 624-629 and 468-498 of Seq ID No 177; 11-19, 32-49, 57-63, 65-71, 80-89, 91-133, 166-181, 183-191, 201-230, 234-257, 264-291, 297-303, 305-314, 316-335, 337-354, 359-366, 368-374, 383-388, 394-405, 408-442, 446-470, 483-490, 499-505, 513-538, 544-555, 557-563, 568-590, 598-608, 617-623, 627-636, 641-647, 667-685, 687-693, 710-723, 733-739, 742-754, 769-815 and 366-388 of Seq ID No 178; 4-16, 30-35, 42-53, 67-76, 82-87, 101-108, 112-130, 132-138, 147-152, 161-183, 187-208, 218-225, 265-281, 295-303, 305-317, 322-334, 338-357, 360-368, 370-383, 387-394, 400-419, 421-430 and 255-336 of Seq ID No 179; 19-27, 36-47, 59-66, 76-83, 101-112, 118-125, 142-147, 162-180, 185-196, 225-240, 246-263, 286-304, 314-319, 327-333, 353-367 and 194-214 of Seq ID No 180; 14-43, 70-76, 83-89, 111-117, 122-128, 136-145, 163-170, 175-182, 210-219, 246-251, 266-279, 325-331, 338-346, 348-354, 356-363, 368-379, 422-428, 431-441, 450-456, 466-473, 509-515, 532-542, 549-556, 576-586, 605-612, 617-623, 636-653, 660-667, 674-686, 698-719, 726-731, 738-745, 762-783, 818-828, 836-843, 856-862, 871-881, 903-910, 917-922, 935-941, 956-968, 976-986, 1013-1024, 1026-1038, 1059-1069, 1071-1081, 1102-1115, 1118-1126, 1131-1145, 1164-1176, 1187-1200, 1202-1208, 1221-1227, 1236-1245, 1248-1267, 1273-1292, 252-287 and 805-844 of Seq ID No 181; 4-18, 21-28, 37-43, 56-70, 101-113, 131-140, 142-150, 162-170, 172-184, 193-204, 209-227, 233-238, 246-264 and 93-168 of Seq ID No 182; 14-20, 44-50, 61-70, 77-96, 99-106, 129-142, 168-181, 187-196, 205-221, 225-241, 277-296 and 257-281 of Seq ID No 183; 18-29, 43-54, 64-76, 78-84, 88-103, 125-149, 159-176, 198-218, 230-242, 256-271, 279-285, 287-293, 300-306, 325-331, 344-351, 357-364, 371-397, 400-414, 419-464, 485-515, 517-526, 529-537, 548-553, 573-580, 584-590, 603-620, 639-661, 676-681, 687-700, 716-761, 772-780, 785-790, 795-803, 823-836, 848-853 and 106-134 of Seq ID No 184; 7-13, 19-42, 44-51, 55-75, 87-97, 99-110, 112-118, 129-135, 141-156, 158-178, 213-220, 230-286, 294-308, 323-338, 345-352, 355-365, 370-392, 394-419, 437-446, 454-460, 474-497, 515-526, 528-546, 569-575 and 128-141 of Seq ID No 185; 12-20, 24-33, 45-70, 73-84, 86-94, 103-116, 118-124, 135-142, 163-170, 176-200, 202-224, 226-234, 237-248, 250-262, 265-287, 296-307, 334-341, 347-356, 361-369, 382-396, 405-415, 418-427, 431-439, 443-449, 452-461, 467-474 and 113-146 of Seq ID No 186; 13-38, 44-50, 52-59, 66-72, 83-94, 103-110, 116-124, 131-137, 158-180, 199-204, 218-233, 241-264, 269-317, 326-342, 350-356 and 70-86 of Seq ID No 187; 29-35, 49-59, 63-84, 86-97, 103-111, 113-126, 130-144, 150-158, 174-198, 221-231, 250-264, 266-273, 291-298, 310-318 and 70-90 of Seq ID No 188; 19-25, 28-52, 60-66, 71-76, 131-142, 149-155, 157-178, 181-213, 218-223, 237-242, 250-257, 260-266, 272-279, 282-290, 321-330, 373-385, 393-407, 441-453, 461-475, 509-521, 529-542, 577-589, 597-610, 643-655, 663-677, 703-718, 729-734, 358-464, 495-570 and 604-685 of Seq ID No 189; 4-29, 51-76, 116-136, 158-173, 179-193, 207-215 and 86-111 of Seq ID No 190; 5-23, 45-70, 79-90, 93-107, 114-122, 142-151 and 18-36 of Seq ID No 191; 9-51, 68-120, 133-149, 158-180, 186-206, 211-220, 222-237, 248-293, 296-310, 317-339 and 248-260 of Seq ID No 192; 14-24, 44-63, 69-98, 108-119, 123-136, 155-161, 164-176, 180-193, 203-208, 215-223, 239-247, 274-281, 283-289, 296-304, 306-313, 315-327, 331-341, 343-353, 357-386, 392-405 and 205-246 of Seq ID No 193; 5-13, 16-23, 36-42, 53-63, 70-83, 96-102 and 14-34 of Seq ID No 194; 4-13, 19-35, 49-56, 59-76, 83-107, 121-134, 144-153, 157-164, 166-186,

194-202, 209-216, 231-253, 257-264 and 98-134 of Seq ID No 195; 16-32, 38-47, 58-68, 78-89, 98-114, 117-123, 132-141, 146-156, 164-170, 179-188, 196-212, 219-230, 232-237, 244-263, 265-274, 278-293, 297-303, 306-326, 339-349, 352-359, 362-367, 373-379, 384-394, 396-406, 423-443, 451-461, 465-484, 490-497, 504-511, 523-533, 537-547, 550-556, 558-566, 573-579, 586-593, 598-609, 615-642, 647-665, 671-686, 693-713, 723-728 and 332-378 of Seq ID No 196; 6-21, 34-44, 58-64, 66-74, 79-87, 114-127, 129-143, 154-162, 174-189, 205-214, 241-262, 266-273, 278-297, 319-324, 328-338, 342-351, 390-398, 409-415, 422-435, 458-464, 471-477, 481-486, 506-531, 534-540, 542-550 and 315-389 of Seq ID No 197; 4-28, 39-45, 52-58, 69-82, 93-115, 122-128, 135-140, 146-163, 177-192, 209-215, 221-232, 271-284, 331-337, 341-352, 360-378, 383-390, 392-401, 409-422, 428-435, 462-470, 474-480, 482-496, 531-539, 541-549, 551-560, 562-569, 576-582, 598-618 and 98-127 of Seq ID No 198; 14-27, 33-47, 61-79, 94-104, 119-133 and 36-60 of Seq ID No 199; 11-22, 29-40, 48-62, 68-73, 96-106, 108-118, 125-149 and 102-126 of Seq ID No 200; 4-11, 45-55, 76-83, 86-102, 105-112, 138-144, 147-153 and 20-48 of Seq ID No 201; 12-20, 28-56, 62-68, 72-82, 93-99, 101-107, 120-133, 135-145, 178-186, 208-232, 279-292 and 36-64 of Seq ID No 202; 6-14, 23-48, 65-82, 92-134, 140-181, 188-219, 228-238, 244-253, 255-261 and 124-145 of Seq ID No 203; 11-25, 31-38, 53-59, 62-71, 89-99, 125-133, 151-157, 182-190, 195-203, 208-215, 219-229, 249-262, 267-275, 287-295, 298-316, 318-325, 328-334, 344-353, 357-363, 371-377, 385-391, 396-415, 425-436, 438-457, 471-485, 538-552, 554-561, 606-625, 630-636, 646-653, 669-679, 695-704, 706-715, 722-747, 763-773 and 714-738 of Seq ID No 204; 10-29, 33-45, 50-60, 70-79, 83-95, 118-124, 136-157, 176-184, 192-205, 207-216, 223-234, 240-246, 258-268, 275-283 and 37-56 of Seq ID No 205; 4-24, 27-38, 46-54, 66-72, 81-97, 112-119, 128-137, 152-157, 173-179, 185-214, 219-225, 227-248, 262-284, 286-295, 301-307 and 117-134 of Seq ID No 206; 26-43, 49-56, 60-71, 74-82, 87-98, 110-116, 131-146, 154-164, 169-178, 183-189, 205-214, 241-246, 255-268, 275-292, 305-314, 316-323, 326-340, 346-363, 397-402, 419-429, 440-446, 452-461, 467-475 and 29-66 of Seq ID No 207; 7-16, 21-39, 48-58, 61-78, 82-89, 109-136, 138-150, 152-176, 182-247, 255-261, 267-332, 336-345, 347-358, 362-368, 371-392, 394-404, 407-472, 490-498, 505-513, 527-544, 554-582, 603-611, 614-620, 632-638 and 500-523 of Seq ID No 208; 24-46, 77-83, 90-97, 99-118, 123-166, 168-177, 204-212, 229-239, 248-262, 273-282, 287-293, 300-319, 321-337, 340-352, 357-366, 391-402, 411-428, 442-450, 464-471, 479-489 and 19-40 of Seq ID No 209; 9-23, 25-34, 53-58, 70-86, 90-97, 99-116, 118-128, 131-141, 185-191, 228-233, 237-253, 255-261, 264-271, 273-280, 302-312, 319-349, 351-359, 362-369, 376-383, 387-394, 398-406, 419-434 and 20-31 of Seq ID No 210; 15-22, 37-43, 71-87, 105-115, 121-127, 135-142, 152-158 and 32-52 of Seq ID No 211; 6-12, 18-29, 37-47, 50-58, 65-83, 85-91, 94-99, 108-123, 142-150, 156-163, 183-193, 215-222, 242-249, 252-258, 261-270, 285-308, 318-326 and 1-95 of Seq ID No 212; 9-61, 65-133, 144-155, 166-173, 175-221, 233-276, 278-313, 329-368 and 210-233 of Seq ID No 213; 11-29, 33-39, 46-51, 65-93, 107-113, 134-143, 147-154, 166-177, 181-188, 214-220, 233-243, 263-269 and 112-128 of Seq ID No 214; 8-46, 110-134, 155-167, 174-183, 188-201, 210-230, 253-258, 267-282, 289-299, 312-319, 322-327, 330-337, 365-381, 389-402, 405-411, 419-425, 439-447, 465-472, 489-512, 525-532, 540-554, 577-589, 591-599, 605-614, 616-624, 633-649 and 503-529 of Seq ID No 215; 34-49, 64-70, 90-118, 124-131, 141-152, 159-165 and 112-128 of Seq ID No 216; 5-15, 26-45, 55-72, 80-85, 93-100, 121-133, 142-148, 154-167, 198-205, 209-215, 241-254, 260-265, 271-279 and 244-270 of Seq ID No 217; 4-36, 38-54, 67-83, 122-153, 159-178, 205-212, 232-242, 244-253, 259-268, 281-288, 298-309, 324-331, 334-370, 372-381, 389-401, 403-429, 441-450, 456-462, 465-471, 473-479, 483-504, 508-518, 537-543, 553-565, 578-584, 592-609, 619-625, 658-667, 669-679, 712-719, 722-729, 737-744, 746-752, 758-765 and 180-226 of Seq ID No 218; 6-17, 23-32, 49-56, 61-67, 76-83, 85-103, 105-111, 120-132, 145-171, 175-185, 191-225, 231-246 and 99-128 of Seq ID No 219; 4-24, 28-48, 52-58, 64-79, 87-100, 104-120, 136-152, 159-166 and 150-163 of Seq ID No 220; 15-27, 65-71, 77-99, 104-121, 128-154, 183-216, 223-229, 234-255, 277-287, 296-308 and 77-97 of Seq ID No 221; 8-18, 44-76, 102-109 and 49-57 of Seq ID No 222; 5-14, 28-40, 42-51, 54-60, 77-83, 89-100, 117-124, 146-172, 176-204, 216-231, 237-244, 267-278, 324-334, 342-348, 396-401, 427-433, 438-450, 452-457, 465-471, 473-481, 491-500, 509-515, 523-544, 550-556, 558-569, 589-595, 606-618, 625-632, 640-649, 665-671, 678-688, 691-698, 717-723, 728-734, 781-789, 800-805, 812-821, 833-868, 873-879, 889-905, 929-939, 988-998, 1046-1061, 1073-1079, 1089-1096, 1115-1124, 1132-1140, 1172-1196, 1220-1226, 1231-1249, 1269-1277, 1287-1301, 1307-1330, 1350-1361, 1369-1378, 1387-1412, 1414-1420, 1422-1439, 1484-1491, 1513-1529, 1552-1561, 1576-1583, 1606-1613, 1617-1640, 1647-1654, 1665-1679, 1686-1698, 1709-1727, 1736-1743, 1750-1757, 1771-1790, 1801-1807, 1817-1823, 1831-1842, 1859-1868, 1870-1882, 1884-1891, 1900-1906, 1909-1914, 1929-1935, 1952-1960, 1974-1988, 2002-2011, 2032-2063, 2071-2081, 2116-2124, 2139-2147, 2149-2159, 2163-2190, 2209-2215, 2245-2253, 2282-2287, 2331-2342, 2360-2370, 2379-2393, 2402-2408, 2414-2421, 2423-2430, 2433-2439, 2442-2450,

2472-2478, 2485-2493, 2495-2503, 2506-2512, 2547-2554, 2558-2564, 2615-2625, 2637-2652, 2692-2698, 2700-2706, 2711-2723, 2731-2740, 2748-2753, 2756-2762, 2765-2772, 2781-2798, 2810-2824, 2844-2852, 2885-2899, 2912-2922, 2937-2944, 2947-2970, 2988-2998, 3016-3025, 3032-3037, 3062-3071, 3129-3148, 3156-3161 and 530-607 of Seq ID No 223; 31-36, 57-62, 79-85, 90-96, 99-112, 120-146, 162-185, 193-203, 208-217, 219-226, 239-253, 283-290, 298-304, 306-321, 340-349, 351-361, 365-372, 386-395, 407-438, 473-486, 537-551, 558-568, 576-594, 598-604 and 75-95 of Seq ID No 224; 14-19, 24-30, 34-42, 45-52, 54-64, 66-82, 95-105, 107-118, 126-163, 171-177, 184-201, 210-215, 260-269, 273-279, 288-304, 321-327, 358-364, 370-375, 380-387, 394-404, 407-413, 421-431, 436-451, 465-474, 504-511, 531-552, 578-587, 614-626, 629-636, 638-671, 691-715, 719-729, 733-745, 752-759, 768-777, 785-792, 794-802, 805-824, 844-854, 867-880, 885-891, 893-902, 907-924, 939-948, 955-964, 966-975, 987-1000, 1012-1017, 1023-1028, 1050-1071, 1083-1098, 1102-1115, 1133-1146, 1170-1183, 1204-1211, 1213-1223, 1262-1311, 1313-1319, 1346-1355, 1366-1371, 1383-1405, 1409-1414 and 776-819 of Seq ID No 225; 12-27, 30-38, 54-61, 64-74, 82-96, 103-110, 117-125, 134-140, 147-158, 185-201, 218-225, 232-253, 265-280, 319-325, 350-362, 366-372, 376-386, 464-483, 485-490, 511-521, 531-537, 542-559, 564-574, 593-609, 613-619, 637-642, 668-677 and 195-214 of Seq ID No 226; 4-21, 59-67, 73-79, 84-91, 141-151, 186-197, 203-214, 222-227, 237-245, 255-260, 281-292, 294-311, 336-344, 346-355, 422-437, 459-466, 484-491 and 77-109 of Seq ID No 227; 10-45, 52-61, 63-70, 74-102, 112-122, 124-132, 164-178, 181-205, 212-240, 246-256 and 226-247 of Seq ID No 228; 38-50, 53-63, 78-87, 89-111, 126-152, 169-176, 179-186, 193-228, 254-267, 275-282, 288-304, 309-318, 325-341, 346-353, 358-367, 384-395, 404-427, 429-435, 456-465, 467-501, 510-521, 523-536, 541-548, 552-560, 563-584, 589-595, 597-620, 625-637, 639-645, 661-666, 712-729, 734-741, 743-750, 775-806, 809-816, 818-840, 842-850 and 693-714 of Seq ID No 229; 5-17, 30-37, 52-75, 77-86, 88-107, 112-135, 151-160, 178-222, 226-246, 263-270, 279-294, 306-314, 327-342, 345-352, 374-381, 389-416, 422-429, 435-449, 453-467, 473-500, 512-522, 524-531, 542-549, 552-560, 565-571, 575-586, 594-600, 613-619, 625-633, 635-641, 647-653, 667-674, 680-699, 711-729, 735-741, 764-775, 781-786, 792-798, 805-813, 817-825, 833-842, 850-855, 860-866, 869-910, 917-930, 949-990 and 533-562 of Seq ID No 230; 7-14, 39-46, 61-74, 83-89, 93-99, 110-121, 136-150, 172-180, 182-200, 207-216, 223-236, 238-251, 265-271, 280-288, 294-309, 320-336, 339-354, 362-377, 383-389, 401-407, 435-441, 446-453, 460-465, 472-487, 499-511, 518-528, 533-540, 557-570, 572-587, 631-637, 643-658, 663-669, 672-678, 681-687, 695-706, 714-728 and 118-139 of Seq ID No 231; 5-19, 24-30, 56-64, 69-79, 93-100, 102-111, 117-123, 125-133, 174-182, 185-199, 205-224, 268-275, 311-336 and 102-125 of Seq ID No 232; 6-35, 39-45, 57-62, 80-85, 92-106, 117-122, 126-171, 214-223, 253-260, 268-273, 285-291, 295-306, 315-320, 325-336, 361-366 and 172-202 of Seq ID No 233; 4-13, 24-37, 45-51, 58-66, 84-92, 112-121, 132-141, 151-171, 175-195, 204-212, 222-240, 262-268, 276-295, 305-336, 338-348, 354-362 and 160-183 of Seq ID No 234; 10-16, 24-35, 41-73, 78-104, 111-121, 124-139, 141-148, 150-164, 196-215, 224-241, 249-282, 299-307, 315-357, 368-378, 393-401 and 345-367 of Seq ID No 235; 4-32, 48-53, 61-67, 84-104, 112-118 and 106-130 of Seq ID No 236; 21-28, 31-36, 65-81, 98-105, 115-121, 123-131, 136-142, 155-161, 177-190 and 201-232 of Seq ID No 237; 4-15, 21-27, 33-39, 42-56, 58-64, 68-82, 84-90, 92-98, 113-122, 146-162, 168-175, 177-189, 191-203, 249-268, 279-285, 287-304, 328-342, 349-358, 371-378, 387-393, 404-413, 419-425, 467-479, 487-498, 513-524, 528-539, 541-565, 572-579, 595-606, 626-635, 637-642 and 612-626 of Seq ID No 238; 7-13, 52-70, 76-82, 97-106, 110-117 and 13-45 of Seq ID No 239; 5-10, 12-48, 59-64, 87-102, 107-128, 131-140, 154-161, 165-171, 173-215 and 54-74 of Seq ID No 240; 4-11, 19-28, 34-40, 74-81, 87-98, 126-147, 163-171, 184-193, 205-213 and 49-124 of Seq ID No 241; 7-14, 23-29, 35-40, 61-67, 99-106, 111-122, 124-133, 135-161, 187-206, 216-229, 236-245, 262-268, 271-280 and 256-273 of Seq ID No 242; 4-13, 17-37, 47-54, 85-99, 105-113, 120-132, 147-166, 180-186, 192-199, 204-216 and 127-144 of Seq ID No 243; 14-27, 29-37, 52-62, 68-76, 89-96, 117-123, 125-131, 137-145, 166-195, 205-212, 214-222, 228-235, 258-264, 271-281, 288-296, 308-324, 332-339, 355-361, 365-371 and 268-293of Seq ID No 244; 4-21, 30-42, 54-60, 78-85, 90-110, 141-147, 160-168, 176-185, 194-206, 218-225, 230-245, 251-261, 287-293, 295-304, 320-326, 334-347, 351-362, 386-402, 413-423, 427-433, 439-453, 456-477, 480-493, 507-513, 526-539, 574-581, 591-598, 600-609, 614-632, 655-665, 685-691, 703-712, 742-747, 757-775, 797-803, 813-819, 823-829, 880-887, 901-906, 930-944, 948-958, 962-968, 971-995, 1002-1009, 1017-1023, 1036-1053, 1069-1081, 1107-1124, 1129-1152, 1178-1195, 1211-1223, 1249-1266, 1271-1288, 1334-1340, 1346-1367, 1-63 and 171-189 of Seq ID No 245; 4-22, 52-63, 70-75, 94-104, 112-125, 133-141, 176-199, 209-216, 244-259, 287-299, 336-352, 366-372, 386-399, 421-436, 444-449, 457-466, 481-487, 506-529, 531-540 and 295-378 of Seq ID No 246; 9-30, 43-49, 58-75, 86-96, 119-131, 138-147, 162-167, 181-201, 208-214 and 16-121 of Seq ID No 247; 4-27, 52-58, 80-90, 92-100, 108-114, 118-143, 169-176, 189-198, 247-261, 281-287, 307-317, 323-329, 352-363, 372-381, 396-

411, 413-426, 429-440, 442-450, 456-461, 468-479 and 1-73 of Seq ID No 248; 4-32, 47-52, 57-63, 71-78, 92-104, 126-142, 153-175 and 145-163 of Seq ID No 249; 17-23, 35-41, 51-70, 73-86, 104-125 and 105-129 of Seq ID No 250; 25-32, 41-50, 75-85, 87-103, 115-122, 138-149, 164-171, 188-210, 212-220, 224-234, 256-273, 288-299, 304-310, 330-336, 357-365, 382-390, 399-405, 414-421, 440-446, 454-461, 480-486, 502-514, 518-540, 543-553, 561-567, 572-580, 582-588, 595-630, 633-651, 672-681, 691-709, 760-767, 813-832, 841-848, 852-866, 873-893, 919-925, 927-933, 940-955, 957-978, 984-997, 1000-1010, 1035-1040, 1044-1051, 1058-1064, 1081-1091, 1097-1124, 1129-1138, 1144-1150, 1158-1165, 1170-1180, 909-936 and 1001-1031 of Seq ID No 251; 4-12, 19-26, 31-41, 49-64, 66-86, 101-117, 119-127, 134-142, 152-161, 163-172, 179-188, 209-218, 234-241, 276-291, 294-300, 307-320, 324-341, 346-356, 373-387, 389-397, 410-416, 418-436, 444-454, 460-472, 481-486, 500-507, 511-535, 541-549, 553-559, 579-586, 602-607, 613-620, 628-640, 654-663, 671-678, 681-691, 709-722, 741-754, 766-774, 778-786, 797-803 and 212-226 of Seq ID No 252; 4-10, 15-27, 34-54, 60-73, 79-88, 101-115, 120-136, 154-162, 167-172, 222-240 and 126-195 of Seq ID No 253; 5-16, 18-25, 29-35, 57-63, 86-91, 107-121, 123-131, 170-179, 185-199, 204-226, 250-255, 262-274, 291-296, 325-347 and 1-38 of Seq ID No 254; 7-19, 22-34, 36-42, 48-54, 60-66, 71-76, 104-110, 118-133, 135-145, 158-164, 167-174, 182-193, 196-204, 217-229, 251-290, 293-299, 309-315 and 288-318 of Seq ID No 255; 43-51, 55-61, 66-73, 80-90, 103-127, 133-142, 174-180, 185-196, 203-210, 229-235, 239-251, 258-266, 272-278, 289-314, 316-326, 340-346, 355-361 and 14-27 of Seq ID No 256; 4-25, 27-33, 35-41, 52-74, 76-89, 99-124, 138-144, 146-159, 167-182, 184-191, 193-206, 211-223, 232-240, 249-257, 270-279, 281-287, 293-310, 322-341, 347-356 and 292-322 of Seq ID No 257; 5-13, 28-38, 43-60, 67-72, 98-116, 122-134, 137-151, 167-174, 177-195, 197-216 and 99-195 of Seq ID No 258; 15-33, 35-42, 48-57, 62-68, 73-91, 107-119, 121-153, 173-194, 205-210, 223-228, 234-241, 243-259, 275-298, 308-315, 327-340, 342-370, 376-391, 398-404, 410-419 and 71-122 of Seq ID No 259; 12-39, 43-64, 87-95, 99-105, 114-126, 128-136, 139-147, 212-225 and 107-141 of Seq ID No 260; 6-33, 40-45, 60-75, 79-86, 121-129, 131-137, 161-167, 172-178, 186-195, 203-212, 236-244, 257-264, 278-294, 306-312, 345-358, 368-381, 386-395, 404-410, 412-418 and 198-270 of Seq ID No 261; 18-31, 34-41, 50-56, 69-83, 99-106, 129-141, 147-153, 159-168, 170-178, 190-198, 200-212, 221-232, 237-255, 261-266, 274-292 and 118-216 of Seq ID No 262; 17-47, 61-67, 87-93, 115-121, 126-132, 140-148, 167-173, 179-186, 214-223, 250-255, 264-272, 282-294, 306-318, 338-353, 358-377, 385-401, 414-420, 433-441, 451-457, 470-480, 505-511, 544-550, 571-581, 600-607, 612-618, 631-648, 655-662, 669-681, 693-714, 721-726, 733-740, 757-778, 813-823, 831-838, 851-857, 866-876, 893-905, 912-917, 930-936, 951-963, 971-981, 1008-1019, 1021-1033, 1035-1041, 1054-1064, 1066-1076, 1097-1110, 1113-1121, 1126-1140, 1159-1171, 1182-1195, 1197-1203, 1216-1222, 1231-1240, 1243-1262, 1268-1287 and 738-828 of Seq ID No 263; 19-28, 40-46, 51-57, 68-74, 81-87, 98-108, 111-121 and 20-36 of Seq ID No 264; 4-17, 19-44, 60-69, 80-86, 110-116 and 33-60 of Seq ID No 265; 8-16, 18-28, 42-50, 53-75, 79-86, 94-99, 122-128, 136-142, 149-163, 166-173, 198-212, 254-272, 288-295, 304-318, 324-329, 343-348, 351-364, 367-383, 389-395, 411-420, 427-436 and 11-56 of Seq ID No 266; 19-25 and 6-24 of Seq ID No 267; 6-39, 59-68 and 44-63 of Seq ID No 268; 5-14, 21-28, 38-53 and 29-41 of Seq ID No 269; 4-13, 31-41, 56-65 and 32-56 of Seq ID No 270; 5-12 and 4-21 of Seq ID No 271; 4-18 and 17-32 of Seq ID No 272; 4-10, 23-33 and 14-30 of Seq ID No 273; 26-34, 44-53 and 35-52 of Seq ID No 274; 1-19 of Seq ID No 275; 4-17, 23-30, 32-37 and 6-23 of Seq ID No 276; 5-33, 40-58, 61-66 and 45-66 of Seq ID No 277; 15-41, 61-67 and 41-65 of Seq ID No 278; 4-12, 16-23, 26-37 and 10-29 of Seq ID No 279; 23-39 and 37-55 of Seq ID No 280; 12-20 and 38-55 of Seq ID No 281; 22-37 and 7-22 of Seq ID No 282; 3-14 of Seq ID No 283; 6-16, 43-65, 71-76 and 17-31 of Seq ID No 284; 4-13, 27-39, 42-69 and 17-32 of Seq ID No 285; 4-12, 26-39 and 10-25 of Seq ID No 286; 2-31 of Seq ID No 287; 6-38, 49-62 and 39-55 of Seq ID No 288; 4-10, 24-30 and 2-19 of Seq ID No 289; 12-17, 25-46 and 15-30 of Seq ID No 290; 4-13 and 2-28 of Seq ID No 291; 30-38 and 17-45 of Seq ID No 292 ; 24-33, 55-61 and 31-61 of Seq ID No 293; 4-26, 34-48 and 15-33 of Seq ID No 294; 9-15 and 1-22 of Seq ID No 295; 4-31 and 14-33 of Seq ID No 296; 5-34, 49-55, 64-82 and 69-83 of Seq ID No 297; 33-45 and 21-39 of Seq ID No 298; 7-14, 24-32, 42-65, 79-86 and 50-64 of Seq ID No 299; 13-27, 33-43, 45-62 and 12-37 of Seq ID No 300; 4-15, 17-32 and 10-26 of Seq ID No 301; 4-9, 11-43, 45-75 and 47-69 of Seq ID No 302; 4-18, 22-37 and 17-34 of Seq ID No 303; 4-14 and 5-24 of Seq ID No 304; 7-33, 35-46 and 1-19 of Seq ID No 305; 13-37, 69-75 and 51-69 of Seq ID No 306; 14-24, 26-34, 37-49, 66-78 and 2-25 of Seq ID No 307; 17-46, 52-57, 59-64 and 54-68 of Seq ID No 308; 4-22 and 13-25 of Seq ID No 309; 8-40, 53-63 and 29-50 of Seq ID No 310; 16-28 and 32-40 of Seq ID No 311; 14-20, 22-28, 39-45 and 2-22 of Seq ID No 312; 4-13 and 12-31 of Seq ID No 313; 15-21 and 2-17 of Seq ID No 314; 4-17 and 20-36 of Seq ID No 315; 4-19 and 9-18 of Seq ID No 316; 4-14 and 3-19 of Seq ID No 317; 4-21, 32-40 and 21-39

of Seq ID No 318; 4-13 and 10-27 of Seq ID No 319; 18-31, 39-47, 75-87, 89-98 and 79-99 of Seq ID No 320; 15-21 and 9-24 of Seq ID No 321; 4-14, 18-27, 30-53, 55-64, 68-74, 81-98 and 22-40 of Seq ID No 322; 7-24, 44-51 and 35-60 of Seq ID No 323; 10-47 and 23-37 of Seq ID No 324; 4-10, 12-46 and 7-22 of Seq ID No 325; 20-27 and 1-13 of Seq ID No 326; 6-19, 41-51 and 9-37 of Seq ID No 327; 4-9, 11-17 and 9-23 of Seq ID No 328; 4-17, 23-38, 46-66, 68-85 and 34-46 of Seq ID No 329; 4-18, 34-59, 75-81 and 61-84 of Seq ID No 330; 6-17 and 7-28 of Seq ID No 331; 4-32, 56-61 and 35-52 of Seq ID No 332; 4-14, 27-71, 74-88, 93-110, 115-120, 124-130, 139-154, 161-172 and 146-171 of Seq ID No 333; 4-21 and 3-15 of Seq ID No 334; 12-17 and 9-26 of Seq ID No 335; 10-21, 45-58 and 51-67 of Seq ID No 336; 59-66, 68-84 and 13-42 of Seq ID No 337; 11-16 and 1-16 of Seq ID No 338; 4-19, 23-37 and 10-30 of Seq ID No 339; 19-27, 35-46, 48-66, 82-88, 99-105, 113-119 and 42-59 of Seq ID No 340; 135-147 of Seq ID No 171; 658-682 of Seq ID No 172; 411-427 and 1226-1246 of Seq ID No 173; 794-817 and 801-824 of Seq ID No 174; 468-492 and 474-495 of Seq ID No 177; 366-388 of Seq ID No 178; 266-291, 287-312 and 308-333 of Seq ID No 179; 197-213 and 195-211 of Seq ID No 180; 252-275, 262-285 and 812-830 of Seq ID No 181; 94-112, 97-120 and 104-128 of Seq ID No 182; 257-281 of Seq ID No 183; 106-134 of Seq ID No 184; 70-86 of Seq ID No 187; 358-383, 378-402, 397-421, 499-524, 520-545, 541-566, 622-646, 641-665 and 660-684 of Seq ID No 189; 248-260 of Seq ID No 192; 15-34 of Seq ID No 194; 112-129 of Seq ID No 195; 333-358 and 353-378 of Seq ID No 196; 316-343, 339-366 and 362-389 of Seq ID No 197; 98-123 and 104-126 of Seq ID No 198; 20-43 and 23-48 of Seq ID No 201; 124-145 of Seq ID No 203; 717-738 of Seq ID No 204; 37-56 of Seq ID No 205; 118-134 of Seq ID No 206; 500-522 of Seq ID No 208; 32-47 of Seq ID No 211; 25-51, 47-73 and 69-95 of Seq ID No 212; 503-529 of Seq ID No 215; 112-128 of Seq ID No 216; 181-199 of Seq ID No 218; 109-121 of Seq ID No 219; 150-163 of Seq ID No 220; 77-97 of Seq ID No 221; 564-586 of Seq ID No 223; 75-94 of Seq ID No 224; 776-798, 784-808 and 794-815 of Seq ID No 225; 196-212, 78-100 and 85-107 of Seq ID No 226; 536-553 of Seq ID No 230; 102-125 of Seq ID No 232; 178-198 of Seq ID No 233; 612-626 of Seq ID No 238; 171-187 of Seq ID No 245; 296-320, 315-339, 334-358 and 353-377 of Seq ID No 246; 47-71 of Seq ID No 247; 1-25, 20-45 and 40-64 of Seq ID No 248; 146-161 of Seq ID No 249; 910-935 and 1007-1030 of Seq ID No 251; 212-226 of Seq ID No 252; 126-152, 148-173 and 169-195 of Seq ID No 253; 288-310 and 293-316 of Seq ID No 255; 293-312 of Seq ID No 257; 154-170 of Seq ID No 258; 72-95, 90-112 and 97-121 of Seq ID No 259; 135-150 and 146-163 of Seq ID No 262; 799-827 of Seq ID No 263; 23-43 and 33-53 of Seq ID No 266; 44-62 of Seq ID No 268; 6-22 of Seq ID No 276; 37-54 of Seq ID No 280; 40-54 of Seq ID No 281; 7-21 of Seq ID No 282; 4-11, 16-34, 48-55, 67-77, 87-106 and 153-183 of Seq ID No 425; 22-40, 49-65, 70-91, 95-109, 111-125, 146-207, 209-216, 219-225, 229-244, 251-270, 274-286, 292-309, 316-329, 335-355, 358-370, 376-388, 392-419, 425-430, 435-441, 448-455, 464-478, 486-515 and 437-465 of Seq ID No 426; 5-19, 25-31, 43-48, 60-79, 88-100, 105-129, 148-171, 187-193, 243-263, 316-322, 334-340, 345-351, 369-378, 381-391, 399-404, 474-483, 502-517, 525-530, 558-568, 579-596, 622-627, 631-638, 644-651, 653-660, 674-680, 687-693, 721-728, 743-753, 760-775, 788-795, 806-813, 821-828, 835-842, 847-859, 868-887 and 300-347 of Seq ID No 427; 5-26, 37-44, 89-97, 112-118, 121-128, 138-154, 157-165, 176-181, 188-198, 205-218, 223-243, 247-253, 260-279 and 76-155 of Seq ID No 428; 4-29, 41-46, 49-68, 82-88, 121-147, 158-164, 187-193, 195-208, 229-236, 244-249, 251-263, 269-275, 307-313, 337-343, 348-381, 392-398, 402-408, 432-438, 85-117 and 194-239 of Seq ID No 429; 5-12, 14-22, 28-34, 40-46, 70-79, 84-129, 152-165, 174-182 and 37-109 of Seq ID No 430; 5-16, 18-52, 54-72, 81-86, 118-126, 136-145, 151-157, 168-180, 209-233, 244-270, 295-302, 315-326, 329-337, 345-352, 364-373, 397-402, 408-418, 424-431, 435-443, 472-480, 483-489, 504-510, 519-527, 549-564, 576-599, 605-637, 641-673 and 91-98 of Seq ID No 431; 23-36, 42-52, 133-140, 151-157, 242-247, 267-277, 295-301, 320-328, 333-339, 345-352, 365-371, 397-403, 415-428, 456-465, 481-487, 489-495, 508-516, 518-527, 585-592, 606-614, 631-637, 643-658, 665-670, 723-728, 737-744, 752-759, 787-793, 835-841, 873-885, 918-928, 938-945, 951-966, 978-988, 1015-1020, 1030-1036, 1044-1052, 1058-1069, 1071-1079, 1081-1088, 1113-1119, 1125-1138, 1141-1147, 1164-1170, 1172-1177, 1190-1200, 1214-1220, 1230-1236, 1239-1245, 1262-1268, 1270-1275, 1288-1298, 1312-1318, 1328-1334, 1337-1343, 1360-1366, 1368-1373, 1386-1396, 1410-1416, 1426-1432, 1435-1441, 1458-1464, 1466-1471, 1484-1494, 1508-1514, 1524-1530, 1533-1539, 1556-1562 and 307-340 of Seq ID No 432; 19-25, 35-41, 44-50, 66-72, 74-79, 92-102, 116-122, 132-138, 141-147, 164-170, 172-177, 190-200, 214-220, 230-236, 239-245, 262-268, 270-275, 288-298, 312-318, 328-334, 337-343, 360-366, 368-373, 386-396, 410-416, 426-432, 435-441, 458-464, 466-478, 504-524, 79-148, 177-246, 275-344 and 373-442 of Seq ID No 433; 7-14, 16-23, 33-39, 46-53, 72-79, 92-115, 123-130, 156-175, 179-187, 214-220, 239-246, 266-274, 302-325, 338-354, 360-370, 375-390, 392-401, 421-428,

430-463 and 29-58 of Seq ID No 434; 4-9, 22-39, 58-65, 72-82, 87-92, 99-104, 107-119, 143-166, 171-177, 194-202, 205-213, 220-228, 231-240, 247-263, 309-315, 317-323, 336-343 and 294-320 of Seq ID No 435; 4-10, 12-18, 24-29, 34-43, 50-65, 70-76, 111-117, 129-138, 152-159, 166-171, 184-195, 200-210, 224-236, 241-251, 274-283, 285-296, 313-319, 332-341, 348-355, 378-386, 410-416, 433-445, 475-482, 523-529, 531-540, 584-596, 626-633, 674-680, 682-688, 738-750, 780-787, 828-834, 836-842, 853-862, 882-887, 893-912 and 604-676 of Seq ID No 436; 15-38, 49-57, 60-99, 103-119, 124-194, 200-206, 215-249, 251-291, 307-313, 315-347, 369-378, 383-390, 393-400, 405-411, 423-435, 440-447, 454-460, 470-486, 490-503, 532-539, 542-549, 551-567, 579-592 and 509-583 of Seq ID No 437; 38-44, 47-88, 95-103, 157-172, 235-240, 250-260, 263-276, 294-300, 312-317, 331-337, 369-391, 412-419, 442-448, 453-463, 490-529, 537-555, 571-580, 600-617, 619-627, 642-648, 682-687, 693-700, 716-722, 738-748, 756-763, 779-789, 796-802, 820-828, 833-840, 846-853, 862-872, 880-887, 894-899, 924-937, 957-963, 1006-1012, 1043-1049, 1063-1069, 1076-1097 and 124-147 of Seq ID No 438; 4-28, 31-49, 60-71, 75-102, 104-114, 134-144, 160-184, 250-257, 277-285, 287-294, 330-338, 345-351, 367-374, 381-388, 393-399, 402-407, 420-426, 443-448, 458-464, 411-436 and 454-488 of Seq ID No 439; 20-27, 45-55, 57-64, 66-77, 98-106, 130-137, 155-165, 167-174, 176-187, 194-203, 208-223, 227-238, 245-251, 257-270, 273-278, 287-299, 330-345, 352-358, 363-385, 392-399, 410-417, 437-443, 467-484, 486-492, 495-500, 504-516, 526-536 and 219-270 of Seq ID No 440; 11-22, 24-31, 46-63, 65-71, 73-88, 95-109, 174-181, 183-201, 204-212, 216-222, 228-233, 241-247 and 142-221 of Seq ID No 441; 8-28, 51-59, 67-84, 93-98, 140-152, 154-162, 183-188 and 91-125 of Seq ID No 442; 10-22, 27-61 and 69-100 of Seq ID No 443; 7-15, 18-26, 94-100, 126-131, 152-165, 219-228, 254-263, 274-292, 297-308, 333-340, 342-352, 354-371, 373-379, 403-410, 420-438, 450-456, 463-470, 489-495, 503-512 and 97-173 of Seq ID No 444; 4-21, 37-43, 49-65, 67-74, 76-90, 113-119, 131-141, 155-173, 175-189, 192-199, 207-221, 247-254, 266-276, 317-322, 337-343, 387-393, 408-428, 439-448, 451-460, 469-479, 482-487, 493-501, 517-523, 533-542 and 480-503 of Seq ID No 445; 11-26, 40-46, 78-86, 93-103, 121-126, 132-138, 166-177, 183-196, 203-212, 214-221, 228-263, 304-311, 323-338, 345-351, 357-363, 379-393, 420-434, 442-448, 518-527, 547-553, 581-591, 602-609, 637-645, 665-674, 687-692, 701-708, 730-739, 796-802, 844-857, 882-888, 903-914, 944-950, 976-983, 1027-1033, 1049-1057, 1066-1072, 1085-1092, 1120-1127, 1137-1144, 1153-1158, 1165-1176, 1181-1187, 1221-1230, 1238-1244, 1269-1274 and 605-632 of Seq ID No 446; 6-47, 57-65, 83-95, 109-121, 138-147, 154-164, 167-177, 194-200, 202-212, 227-234, 240-253, 260-267, 283-291, 320-329, 340-347, 356-364, 412-422, 430-436, 441-459, 465-475, 478-486, 498-507 and 59-84 of Seq ID No 447; 10-21, 58-83, 88-97, 120-126 and 21-51 of Seq ID No 448; 5-39, 56-62, 76-88, 90-114, 138-162, 170-195, 202-221, 228-250, 264-270, 304-355, 374-387, 391-416, 462-471, 526-546, 554-561, 574-579, 639-645, 651-660, 674-682, 689-694 and 666-697 of Seq ID No 449; 6-30, 36-42, 143-157, 176-197, 202-209, 216-233, 241-246, 275-287, 292-299, 315-325, 343-350, 375-380, 397-403, 411-420, 422-434, 441-448, 467-474, 477-499, 555-568, 591-597, 601-609, 623-644, 667-688, 692-698, 703-718, 736-747, 757-766, 782-791, 795-801, 832-840, 859-865 and 226-269 of Seq ID No 450; 6-23, 43-51, 61-67, 73-82, 91-97, 123-130, 149-158, 164-175, 228-234, 240-246, 248-255, 262-272, 326-332, 340-347, 365-371, 377-388, 409-419, 425-431, 438-445, 449-457, 464-470, 496-507, 559-568, 575-581, 603-608, 617-623, 633-639, 648-654, 659-670, 695-701, 734-752, 806-814, 816-829, 861-868, 891-899, 904-909, 937-945, 947-960, 978-983, 992-999, 1022-1031, 1068-1076, 1078-1091, 1109-1114, 1123-1130, 1153-1162, 1199-1207, 1209-1222, 1254-1261, 1284-1293, 1330-1338, 1340-1353, 1371-1376, 1385-1392, 1415-1421, 1433-1438, 1460-1465, 1470-1492 and 1422-1458 of Seq ID No 451; 82-94, 111-118, 125-131, 206-212, 261-266, 310-320, 328-338, 345-351, 353-360, 414-420, 424-434, 440-447, 451-500, 506-516, 548-561, 566-572, 584-591, 601-622, 630-636, 650-659, 661-667, 674-699, 703-711, 717-729, 736-744, 752-759, 765-771, 813-822, 826-842, 852-868, 870-877, 881-895, 897-906, 913-922 and 602-671 of Seq ID No 452; 12-18, 20-25, 43-54, 56-65, 73-79, 82-88, 99-111, 136-142, 153-169, 171-183, 195-223, 229-248, 255-260, 272-277, 281-292, 298-319, 322-329, 332-351, 363-379, 381-389 and 275-304 of Seq ID No 453; 4-9, 34-48, 65-77, 101-106, 111-131, 138-153, 186-191, 230-250 and 148-219 of Seq ID No 454; 4-23, 30-35, 42-53, 67-76, 82-87, 101-108, 112-130, 132-138, 147-152, 161-183, 187-208, 218-225, 265-283, 295-303, 306-317, 322-334, 338-357, 360-368, 370-383, 387-398, 400-419, 421-430, 104-182 and 240-304 of Seq ID No 455; 4-12, 63-69, 94-102, 146-164, 166-173, 175-181, 193-207, 263-281, 286-295, 301-306, 330-343, 369-378, 382-388, 414-420, 422-430, 438-454, 456-462, 472-531, 543-560, 581-591, 596-605, 614-623, 626-635, 656-662, 669-676, 683-690, 693-698, 705-711, 728-736, 752-764 and 69-102 of Seq ID No 456; 6-12, 43-53, 141-147, 164-179, 185-195, 197-206, 227-235, 237-271, 288-305, 308-317, 335-341, 351-357, 365-376, 386-395, 397-416, 422-447 and 11-35 of Seq ID No 457; 16-24, 50-65, 73-84, 88-99, 114-124, 130-146, 181-187, 193-203, 214-220, 236-247, 250-258, 287-297 and 50-113 of Seq ID No 458; 4-25, 50-55, 76-82, 117-123, 131-

137, 139-148, 157-166, 239-245, 253-258, 266-275, 277-292, 300-306, 51-83 and 93-161 of Seq ID No 459; 6-22, 34-43, 51-86, 93-100, 110-116, 150-161, 164-171, 180-187, 197-218 and 168-237 of Seq ID No 460; 4-27, 55-60, 74-82 and 10-46 of Seq ID No 461; 6-19, 25-31, 43-49, 60-79, 88-100, 105-129, 148-161, 164-171, 187-193, 243-263, 316-322, 334-340, 369-378, 381-391, 398-404, 460-466, 474-483, 502-509, 511-517, 525-530, 558-567, 579-596, 622-627, 631-638, 641-651, 653-659, 674-680, 687-693, 710-716, 720-727, 743-753, 760-775, 788-795, 806-813, 821-828, 836-842, 847-860, 865-880 and 258-377 of Seq ID No 462; 4-11, 25-64, 71-79, 88-94, 107-120, 123-132, 167-188, 231-237, 240-246, 261-267, 306-311, 330-342, 351-358, 389-395, 406-418, 429-434, 439-448, 483-501, 511-520 and 71-143 of Seq ID No 463; 4-18, 22-27, 53-64, 94-100, 121-127, 133-139, 155-164, 177-182, 187-196, 206-218, 224-242, 248-253, 258-277 and 184-253 of Seq ID No 464; 10-17, 56-67, 72-82, 94-99, 106-113, 166-173, 229-235, 243-283, 295-301, 313-321, 326-331, 342-348, 396-414, 423-435, 446-452, 454-462, 496-502, 511-534, 543-556, 563-570, 586-593, 616-626, 638-645, 653-662, 679-696, 731-737, 766-774, 776-782, 790-796, 810-817, 825-835, 837-846 and 540-615 of Seq ID No 465; 13-24, 30-36, 73-81, 89-95, 109-115, 117-143, 161-173, 179-189, 226-244, 251-261, 275-281, 298-305, 307-315, 323-328, 364-374, 69-186 and 264-354 of Seq ID No 466; 19-25 and 6-22 of Seq ID No 467; 6-39, 59-68 and 43-62 of Seq ID No 468; 6-14, 22-32 and 1-27 of Seq ID No 469; 4-41 and 28-40 of Seq ID No 470; 8-14 and 4-19 of Seq ID No 471; 4-10, 12-22, 30-35 and 6-33 of Seq ID No 472; 4-16, 24-33 and 37-54 of Seq ID No 473; 2-23 of Seq ID No 474; 4-21, 27-33, 36-41 and 14-34 of Seq ID No 475; 4-14, 24-30, 37-42, 57-78, 83-89, 94-103, 113-131 and 100-122 of Seq ID No 476, and fragments comprising at least 6, preferably more than 8, especially more than 10 aa and preferably not more than 70, 50, 40, 20, 15, 11 aa of said sequences. All these fragments individually and each independently form a preferred selected aspect of the present invention.

All linear hyperimmune serum reactive fragments of a particular antigen may be identified by analysing the entire sequence of the protein antigen by a set of peptides overlapping by 1 amino acid with a length of at least 10 amino acids. Subsequently, non-linear epitopes can be identified by analysis of the protein antigen with hyperimmune sera using the expressed full-length protein or domain polypeptides thereof. Assuming that a distinct domain of a protein is sufficient to form the 3D structure independent from the native protein, the analysis of the respective recombinant or synthetically produced domain polypeptide with hyperimmune serum would allow the identification of conformational epitopes within the individual domains of multi-domain proteins. For those antigens where a domain possesses linear as well as conformational epitopes, competition experiments with peptides corresponding to the linear epitopes may be used to confirm the presence of conformational epitopes.

It will be appreciated that the invention also relates to, among others, nucleic acid molecules encoding the aforementioned fragments, nucleic acid molecules that hybridise to nucleic acid molecules encoding the fragments, particularly those that hybridise under stringent conditions, and nucleic acid molecules, such as PCR primers, for amplifying nucleic acid molecules that encode the fragments. In these regards, preferred nucleic acid molecules are those that correspond to the preferred fragments, as discussed above.

The present invention also relates to vectors which comprise a nucleic acid molecule or nucleic acid molecules of the present invention, host cells which are genetically engineered with vectors of the invention and the production of hyperimmune serum reactive antigens and fragments thereof by recombinant techniques.

A great variety of expression vectors can be used to express a hyperimmune serum reactive antigen or fragment thereof according to the present invention. Generally, any vector suitable to maintain, propagate or express nucleic acids to express a polypeptide in a host may be used for expression in this regard. In accordance with this aspect of the invention the vector may be, for example, a plasmid vector, a single or double-stranded phage vector, a single or double-stranded RNA or DNA viral vector. Starting plasmids disclosed herein are either commercially available, publicly available, or can be constructed from available plasmids by routine application of well-known, published procedures. Preferred among vectors, in certain respects, are those for expression of nucleic acid molecules and hyperimmune serum

reactive antigens or fragments thereof of the present invention. Nucleic acid constructs in host cells can be used in a conventional manner to produce the gene product encoded by the recombinant sequence. Alternatively, the hyperimmune serum reactive antigens and fragments thereof of the invention can be synthetically produced by conventional peptide synthesizers. Mature proteins can be expressed in mammalian cells, yeast, bacteria, or other cells under the control of appropriate promoters. Cell-free translation systems can also be employed to produce such proteins using RNAs derived from the DNA construct of the present invention.

Host cells can be genetically engineered to incorporate nucleic acid molecules and express nucleic acid molecules of the present invention. Representative examples of appropriate hosts include bacterial cells, such as enterococci, staphylococci, E. coli, Streptomyces and Bacillus subtillis cells; fungal cells, such as yeast cells and Aspergillus cells; insect cells such as Drosophila S2 and Spodoptera Sf9 cells; animal cells such as CHO, COS, Hela, C127, 3T3, BHK, 293 and Bowes melanoma cells; and plant cells.

The invention also provides a process for producing a E. faecalis hyperimmune serum reactive antigen and a fragment thereof comprising expressing from the host cell a hyperimmune serum reactive antigen or fragment thereof encoded by the nucleic acid molecules provided by the present invention. The invention further provides a process for producing a cell, which expresses a E. faecalis hyperimmune serum reactive antigen or a fragment thereof comprising transforming or transfecting a suitable host cell with the vector according to the present invention such that the transformed or transfected cell expresses the polypeptide encoded by the nucleic acid contained in the vector.

The polypeptide may be expressed in a modified form, such as a fusion protein, and may include not only secretion signals but also additional heterologous functional regions. Thus, for instance, a region of additional amino acids, particularly charged amino acids, may be added to the N- or C-terminus of the polypeptide to improve stability and persistence in the host cell, during purification or during subsequent handling and storage. Also, regions may be added to the polypeptide to facilitate purification. Such regions may be removed prior to final preparation of the polypeptide. The addition of peptide moieties to polypeptides to engender secretion or excretion, to improve stability or to facilitate purification, among others, are familiar and routine techniques in the art. A preferred fusion protein comprises a heterologous region from immunoglobulin that is useful to solubilize or purify polypeptides. For example, EP-A-O 464 533 (Canadian counterpart 2045869) discloses fusion proteins comprising various portions of constant region of immunoglobin molecules together with another protein or part thereof. In drug discovery, for example, proteins have been fused with antibody Fc portions for the purpose of high-throughout screening assays to identify antagonists. See for example, {Bennett, D. et al., 1995} and {Johanson, K. et al., 1995}.

The E. faecalis hyperimmune serum reactive antigen or a fragment thereof can be recovered and purified from recombinant cell cultures by well-known methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, hydroxylapatite chromatography and lectin chromatography.

The hyperimmune serum reactive antigens and fragments thereof according to the present invention can be produced by chemical synthesis as well as by biotechnological means. The latter comprise the transfection or transformation of a host cell with a vector containing a nucleic acid according to the present invention and the cultivation of the transfected or transformed host cell under conditions which are known to the ones skilled in the art. The production method may also comprise a purification step in order to purify or isolate the polypeptide to be manufactured. In a preferred embodiment the vector is a vector according to the present invention.

The hyperimmune serum reactive antigens and fragments thereof according to the present invention may be used for the detection of the organism or organisms in a sample containing these organisms or polypeptides derived thereof. Preferably such detection is for diagnosis, more preferable for the diagnosis of a disease, most preferably for the diagnosis of a diseases related or linked to the presence or abundance of Gram-positive bacteria, especially bacteria selected from the group comprising enterococci, staphylococci and lactococci. More preferably, the microorganisms are selected from the group comprising *Streptococcus agalactiae*, *Streptococcus pneumoniae* and *Streptococcus mutans*, especially the microorganism is *Enterococcus faecalis*.

The present invention also relates to diagnostic assays such as quantitative and diagnostic assays for detecting levels of the hyperimmune serum reactive antigens and fragments thereof of the present invention in cells and tissues, including determination of normal and abnormal levels. Thus, for instance, a diagnostic assay in accordance with the invention for detecting over-expression of the polypeptide compared to normal control tissue samples may be used to detect the presence of an infection, for example, and to identify the infecting organism. Assay techniques that can be used to determine levels of a polypeptide, in a sample derived from a host are well-known to those of skill in the art. Such assay methods include radioimmunoassays, competitive-binding assays, Western Blot analysis and ELISA assays. Among these, ELISAs frequently are preferred. An ELISA assay initially comprises preparing an antibody specific to the polypeptide, preferably a monoclonal antibody. In addition, a reporter antibody generally is prepared which binds to the monoclonal antibody. The reporter antibody is attached to a detectable reagent such as radioactive, fluorescent or enzymatic reagent, such as horseradish peroxidase enzyme.

The hyperimmune serum reactive antigens and fragments thereof according to the present invention may also be used for the purpose of or in connection with an array. More particularly, at least one of the hyperimmune serum reactive antigens and fragments thereof according to the present invention may be immobilized on a support. Said support typically comprises a variety of hyperimmune serum reactive antigens and fragments thereof whereby the variety may be created by using one or several of the hyperimmune serum reactive antigens and fragments thereof according to the present invention and/or hyperimmune serum reactive antigens and fragments thereof being different. The characterizing feature of such array as well as of any array in general is the fact that at a distinct or predefined region or position on said support or a surface thereof, a distinct polypeptide is immobilized. Because of this any activity at a distinct position or region of an array can be correlated with a specific polypeptide. The number of different hyperimmune serum reactive antigens and fragments thereof immobilized on a support may range from as little as 10 to several 1000 different hyperimmune serum reactive antigens and fragments thereof. The density of hyperimmune serum reactive antigens and fragments thereof per cm² is in a preferred embodiment as little as 10 peptides/polypeptides per cm² to at least 400 different peptides/polypeptides per cm² and more particularly at least 1000 different hyperimmune serum reactive antigens and fragments thereof per cm².

The manufacture of such arrays is known to the one skilled in the art and, for example, described in US patent 5,744,309. The array preferably comprises a planar, porous or non-porous solid support having at least a first surface. The hyperimmune serum reactive antigens and fragments thereof as disclosed herein, are immobilized on said surface. Preferred support materials are, among others, glass or cellulose. It is also within the present invention that the array is used for any of the diagnostic applications described herein. Apart from the hyperimmune serum reactive antigens and fragments thereof according to the present invention also the nucleic acid molecules according to the present invention may be used for the generation of an array as described above. This applies as well to an array made of antibodies, preferably monoclonal antibodies as, among others, described herein.

In a further aspect the present invention relates to an antibody directed to any of the hyperimmune serum reactive antigens and fragments thereof, derivatives or fragments thereof according to the present

invention. The present invention includes, for example, monoclonal and polyclonal antibodies, chimeric, single chain, and humanized antibodies, as well as Fab fragments, or the product of a Fab expression library. It is within the present invention that the antibody may be chimeric, i. e. that different parts thereof stem from different species or at least the respective sequences are taken from different species.

Antibodies generated against the hyperimmune serum reactive antigens and fragments thereof corresponding to a sequence of the present invention can be obtained by direct injection of the hyperimmune serum reactive antigens and fragments thereof into an animal or by administering the hyperimmune serum reactive antigens and fragments thereof to an animal, preferably a non-human. The antibody so obtained will then bind the hyperimmune serum reactive antigens and fragments thereof itself. In this manner, even a sequence encoding only a fragment of a hyperimmune serum reactive antigen and fragments thereof can be used to generate antibodies binding the whole native hyperimmune serum reactive antigen and fragments thereof. Such antibodies can then be used to isolate the hyperimmune serum reactive antigens and fragments thereof from tissue expressing those hyperimmune serum reactive antigens and fragments thereof.

For preparation of monoclonal antibodies, any technique known in the art which provides antibodies produced by continuous cell line cultures can be used (as described originally in {Kohler, G. et al., 1975}).

Techniques described for the production of single chain antibodies (U.S. Patent No. 4,946,778) can be adapted to produce single chain antibodies to immunogenic hyperimmune serum reactive antigens and fragments thereof according to this invention. Also, transgenic mice, or other organisms such as other mammals, may be used to express humanized antibodies to immunogenic hyperimmune serum reactive antigens and fragments thereof according to this invention.

Alternatively, phage display technology or ribosomal display could be utilized to select antibody genes with binding activities towards the hyperimmune serum reactive antigens and fragments thereof either from repertoires of PCR amplified v-genes of lymphocytes from humans screened for possessing respective target antigens or from naïve libraries {McCafferty, J. et al., 1990}; {Marks, J. et al., 1992}. The affinity of these antibodies can also be improved by chain shuffling {Clackson, T. et al., 1991}.

If two antigen binding domains are present, each domain may be directed against a different epitope – termed 'bispecific' antibodies.

The above-described antibodies may be employed to isolate or to identify clones expressing the hyperimmune serum reactive antigens and fragments thereof or purify the hyperimmune serum reactive antigens and fragments thereof of the present invention by attachment of the antibody to a solid support for isolation and/or purification by affinity chromatography.

Thus, among others, antibodies against the hyperimmune serum reactive antigens and fragments thereof of the present invention may be employed to inhibit and/or treat infections, particularly bacterial infections and especially infections arising from E. faecalis.

Hyperimmune serum reactive antigens and fragments thereof include antigenically, epitopically or immunologically equivalent derivatives which form a particular aspect of this invention. The term "antigenically equivalent derivative" as used herein encompasses a hyperimmune serum reactive antigen and fragments thereof or its equivalent which will be specifically recognized by certain antibodies which, when raised to the protein or hyperimmune serum reactive antigen and fragments thereof according to the present invention, interfere with the interaction between pathogen and mammalian host. The term "immunologically equivalent derivative" as used herein encompasses a peptide or its equivalent which when used in a suitable formulation to raise antibodies in a vertebrate, the antibodies act to interfere with the interaction between pathogen and mammalian host.

The hyperimmune serum reactive antigens and fragments thereof, such as an antigenically or immunologically equivalent derivative or a fusion protein thereof can be used as an antigen to immunize a mouse or other animal such as a rat or chicken. The fusion protein may provide stability to the hyperimmune serum reactive antigens and fragments thereof. The antigen may be associated, for example by conjugation, with an immunogenic carrier protein, for example bovine serum albumin (BSA) or keyhole limpet haemocyanin (KLH). Alternatively, an antigenic peptide comprising multiple copies of the protein or hyperimmune serum reactive antigen and fragments thereof, or an antigenically or immunologically equivalent hyperimmune serum reactive antigen and fragments thereof, may be sufficiently antigenic to improve immunogenicity so as to obviate the use of a carrier.

Preferably the antibody or derivative thereof is modified to make it less immunogenic in the individual. For example, if the individual is human the antibody may most preferably be "humanized", wherein the complimentarity determining region(s) of the hybridoma-derived antibody has been transplanted into a human monoclonal antibody, for example as described in {Jones, P. et al., 1986} or {Tempest, P. et al., 1991}.

The use of a polynucleotide of the invention in genetic immunization will preferably employ a suitable delivery method such as direct injection of plasmid DNA into muscle, delivery of DNA complexed with specific protein carriers, coprecipitation of DNA with calcium phosphate, encapsulation of DNA in various forms of liposomes, particle bombardment {Tang, D. et al., 1992}; {Eisenbraun, M. et al., 1993} and *in vivo* infection using cloned retroviral vectors {Seeger, C. et al., 1984}.

In a further aspect the present invention relates to a peptide binding to any of the hyperimmune serum reactive antigens and fragments thereof according to the present invention, and a method for the manufacture of such peptides whereby the method is characterized by the use of the hyperimmune serum reactive antigens and fragments thereof according to the present invention and the basic steps are known to the one skilled in the art.

Such peptides may be generated by using methods according to the state of the art such as phage display or ribosome display. In case of phage display, basically a library of peptides is generated, in form of phages, and this kind of library is contacted with the target molecule, in the present case a hyperimmune serum reactive antigen and fragments thereof according to the present invention. Those peptides binding to the target molecule are subsequently removed, preferably as a complex with the target molecule, from the respective reaction. It is known to the one skilled in the art that the binding characteristics, at least to a certain extent, depend on the particularly realized experimental set-up such as the salt concentration and the like. After separating those peptides binding to the target molecule with a higher affinity or a bigger force, from the non-binding members of the library, and optionally also after removal of the target molecule from the complex of target molecule and peptide, the respective peptide(s) may subsequently be characterised. Prior to the characterisation optionally an amplification step is realized such as, e. g. by propagating the peptide encoding phages. The characterisation preferably comprises the sequencing of the target binding peptides. Basically, the peptides are not limited in their lengths, however, preferably peptides having a lengths from about 8 to 20 amino acids are preferably obtained in the respective methods. The size of the libraries may be about 10^2 to 10^{18}, preferably 10^8 to 10^{15} different peptides, however, is not limited thereto.

A particular form of target binding hyperimmune serum reactive antigens and fragments thereof are the so-called "anticalines" which are, among others, described in German patent application DE 197 42 706.

In a further aspect the present invention relates to functional nucleic acids interacting with any of the hyperimmune serum reactive antigens and fragments thereof according to the present invention, and a method for the manufacture of such functional nucleic acids whereby the method is characterized by the

use of the hyperimmune serum reactive antigens and fragments thereof according to the present invention and the basic steps are known to the one skilled in the art. The functional nucleic acids are preferably aptamers and spiegelmers. .

Aptamers are D-nucleic acids which are either single stranded or double stranded and which specifically interact with a target molecule. The manufacture or selection of aptamers is, e. g., described in European patent EP 0 533 838. Basically the following steps are realized. First, a mixture of nucleic acids, i. e. potential aptamers, is provided whereby each nucleic acid typically comprises a segment of several, preferably at least eight subsequent randomised nucleotides. This mixture is subsequently contacted with the target molecule whereby the nucleic acid(s) bind to the target molecule, such as based on an increased affinity towards the target or with a bigger force thereto, compared to the candidate mixture. The binding nucleic acid(s) are/is subsequently separated from the remainder of the mixture. Optionally, the thus obtained nucleic acid(s) is amplified using, e.g. polymerase chain reaction. These steps may be repeated several times giving at the end a mixture having an increased ratio of nucleic acids specifically binding to the target from which the final binding nucleic acid is then optionally selected. These specifically binding nucleic acid(s) are referred to as aptamers. It is obvious that at any stage of the method for the generation or identification of the aptamers samples of the mixture of individual nucleic acids may be taken to determine the sequence thereof using standard techniques. It is within the present invention that the aptamers may be stabilized such as, e. g., by introducing defined chemical groups which are known to the one skilled in the art of generating aptamers. Such modification may for example reside in the introduction of an amino group at the 2'-position of the sugar moiety of the nucleotides. Aptamers are currently used as therapeutical agents. However, it is also within the present invention that the thus selected or generated aptamers may be used for target validation and/or as lead substance for the development of medicaments, preferably of medicaments based on small molecules. This is actually done by a competition assay whereby the specific interaction between the target molecule and the aptamer is inhibited by a candidate drug whereby upon replacement of the aptamer from the complex of target and aptamer it may be assumed that the respective drug candidate allows a specific inhibition of the interaction between target and aptamer, and if the interaction is specific, said candidate drug will, at least in principle, be suitable to block the target and thus decrease its biological availability or activity in a respective system comprising such target. The thus obtained small molecule may then be subject to further derivatisation and modification to optimise its physical, chemical, biological and/or medical characteristics such as toxicity, specificity, biodegradability and bioavailability.

Spiegelmers and their generation or manufacture is based on a similar principle. The manufacture of spiegelmers is described in international patent application WO 98/08856. Spiegelmers are L-nucleic acids, which means that they are composed of L-nucleotides rather than D-nucleotides as aptamers are. Spiegelmers are characterized by the fact that they have a very high stability in biological systems and, comparable to aptamers, specifically interact with the target molecule against which they are directed. In the process of generating spiegelmers, a heterogeneous population of D-nucleic acids is created and this population is contacted with the optical antipode of the target molecule, in the present case for example with the D-enantiomer of the naturally occurring L-enantiomer of the hyperimmune serum reactive antigens and fragments thereof according to the present invention. Subsequently, those D-nucleic acids are separated which do not interact with the optical antipode of the target molecule. But those D-nucleic acids interacting with the optical antipode of the target molecule are separated, optionally identified and/or sequenced and subsequently the corresponding L-nucleic acids are synthesized based on the nucleic acid sequence information obtained from the D-nucleic acids. These L-nucleic acids which are identical in terms of sequence with the aforementioned D-nucleic acids interacting with the optical antipode of the target molecule, will specifically interact with the naturally occurring target molecule rather than with the optical antipode thereof. Similar to the method for the generation of aptamers it is also possible to repeat the various steps several times and thus to enrich those nucleic acids specifically interacting with the optical antipode of the target molecule.

In a further aspect the present invention relates to functional nucleic acids interacting with any of the nucleic acid molecules according to the present invention, and a method for the manufacture of such functional nucleic acids whereby the method is characterized by the use of the nucleic acid molecules and their respective sequences according to the present invention and the basic steps are known to the one skilled in the art. The functional nucleic acids are preferably ribozymes, antisense oligonucleotides and siRNA.

Ribozymes are catalytically active nucleic acids which preferably consist of RNA which basically comprises two moieties. The first moiety shows a catalytic activity whereas the second moiety is responsible for the specific interaction with the target nucleic acid, in the present case the nucleic acid coding for the hyperimmune serum reactive antigens and fragments thereof according to the present invention. Upon interaction between the target nucleic acid and the second moiety of the ribozyme, typically by hybridisation and Watson-Crick base pairing of essentially complementary stretches of bases on the two hybridising strands, the catalytically active moiety may become active which means that it catalyses, either intramolecularly or intermolecularly, the target nucleic acid in case the catalytic activity of the ribozyme is a phosphodiesterase activity. Subsequently, there may be a further degradation of the target nucleic acid which in the end results in the degradation of the target nucleic acid as well as the protein derived from the said target nucleic acid. Ribozymes, their use and design principles are known to the one skilled in the art, and, for example described in {Doherty, E. et al., 2001} and {Lewin, A. et al., 2001}.

The activity and design of antisense oligonucleotides for the manufacture of a medicament and as a diagnostic agent, respectively, is based on a similar mode of action. Basically, antisense oligonucleotides hybridise based on base complementarity, with a target RNA, preferably with a mRNA, thereby activating RNase H. RNase H is activated by both phosphodiester and phosphorothioate-coupled DNA. Phosphodiester-coupled DNA, however, is rapidly degraded by cellular nucleases with the exception of phosphorothioate-coupled DNA. These resistant, non-naturally occurring DNA derivatives do not inhibit RNase H upon hybridisation with RNA. In other words, antisense polynucleotides are only effective as DNA RNA hybride complexes. Examples for this kind of antisense oligonucleotides are described, among others, in US-patent US 5,849,902 and US 5,989,912. In other words, based on the nucleic acid sequence of the target molecule which in the present case are the nucleic acid molecules for the hyperimmune serum reactive antigens and fragments thereof according to the present invention, either from the target protein from which a respective nucleic acid sequence may in principle be deduced, or by knowing the nucleic acid sequence as such, particularly the mRNA, suitable antisense oligonucleotides may be designed base on the principle of base complementarity.

Particularly preferred are antisense-oligonucleotides which have a short stretch of phosphorothioate DNA (3 to 9 bases). A minimum of 3 DNA bases is required for activation of bacterial RNase H and a minimum of 5 bases is required for mammalian RNase H activation. In these chimeric oligonucleotides there is a central region that forms a substrate for RNase H that is flanked by hybridising "arms" comprised of modified nucleotides that do not form substrates for RNase H. The hybridising arms of the chimeric oligonucleotides may be modified such as by 2'-O-methyl or 2'-fluoro. Alternative approaches used methylphosphonate or phosphoramidate linkages in said arms. Further embodiments of the antisense oligonucleotide useful in the practice of the present invention are P-methoxyoligonucleotides, partial P-methoxyoligodeoxyribonucleotides or P-methoxyoligonucleotides.

Of particular relevance and usefulness for the present invention are those antisense oligonucleotides as more particularly described in the above two mentioned US patents. These oligonucleotides contain no naturally occurring 5'→3'-linked nucleotides. Rather the oligonucleotides have two types of nucleotides: 2'-deoxyphosphorothioate, which activate RNase H, and 2'-modified nucleotides, which do not. The linkages between the 2'-modified nucleotides can be phosphodiesters, phosphorothioate or P-ethoxyphosphodiester. Activation of RNase H is accomplished by a contiguous RNase H-activating

region, which contains between 3 and 5 2'-deoxyphosphorothioate nucleotides to activate bacterial RNase H and between 5 and 10 2'- deoxyphosphorothioate nucleotides to activate eucaryotic and, particularly, mammalian RNase H. Protection from degradation is accomplished by making the 5' and 3' terminal bases highly nuclease resistant and, optionally, by placing a 3' terminal blocking group.

More particularly, the antisense oligonucleotide comprises a 5' terminus and a 3' terminus; and from position 11 to 59 5'→3'-linked nucleotides independently selected from the group consisting of 2'-modified phosphodiester nucleotides and 2'-modified P-alkyloxyphosphotriester nucleotides; and wherein the 5'-terminal nucleoside is attached to an RNase H-activating region of between three and ten contiguous phosphorothioate-linked deoxyribonucleotides, and wherein the 3'-terminus of said oligonucleotide is selected from the group consisting of an inverted deoxyribonucleotide, a contiguous stretch of one to three phosphorothioate 2'-modified ribonucleotides, a biotin group and a P-alkyloxyphosphotriester nucleotide.

Also an antisense oligonucleotide may be used wherein not the 5' terminal nucleoside is attached to an RNase H-activating region but the 3'-terminal nucleoside as specified above. Also, the 5' terminus is selected from the particular group rather than the 3' terminus of said oligonucleotide.

The nucleic acids as well as the hyperimmune serum reactive antigens and fragments thereof according to the present invention may be used as or for the manufacture of pharmaceutical compositions, especially vaccines. Preferably such pharmaceutical composition, preferably vaccine is for the prevention or treatment of diseases caused by, related to or associated with E. faecalis. In so far another aspect of the invention relates to a method for inducing an immunological response in an individual, particularly a mammal, which comprises inoculating the individual with the hyperimmune serum reactive antigens and fragments thereof of the invention, or a fragment or variant thereof, adequate to produce antibodies to protect said individual from infection, particularly enterococcal infection and most particularly E. faecalis infections.

Yet another aspect of the invention relates to a method of inducing an immunological response in an individual which comprises, through gene therapy or otherwise, delivering a nucleic acid functionally encoding hyperimmune serum reactive antigens and fragments thereof, or a fragment or a variant thereof, for expressing the hyperimmune serum reactive antigens and fragments thereof, or a fragment or a variant thereof in vivo in order to induce an immunological response to produce antibodies or a cell mediated T cell response, either cytokine-producing T cells or cytotoxic T cells, to protect said individual from disease, whether that disease is already established within the individual or not. One way of administering the gene is by accelerating it into the desired cells as a coating on particles or otherwise.

A further aspect of the invention relates to an immunological composition which, when introduced into a host capable of having induced within it an immunological response, induces an immunological response in such host, wherein the composition comprises recombinant DNA which codes for and expresses an antigen of the hyperimmune serum reactive antigens and fragments thereof of the present invention. The immunological response may be used therapeutically or prophylactically and may take the form of antibody immunity or cellular immunity such as that arising from CTL or CD4+ T cells.

The hyperimmune serum reactive antigens and fragments thereof of the invention or a fragment thereof may be fused with a co-protein which may not by itself produce antibodies, but is capable of stabilizing the first protein and producing a fused protein which will have immunogenic and protective properties. This fused recombinant protein preferably further comprises an antigenic co-protein, such as Glutathione-S-transferase (GST) or beta-galactosidase, relatively large co-proteins which solubilise the protein and facilitate production and purification thereof. Moreover, the co-protein may act as an adjuvant in the sense of providing a generalized stimulation of the immune system. The co-protein may be attached to either the amino or carboxy terminus of the first protein.

Also, provided by this invention are methods using the described nucleic acid molecule or particular fragments thereof in such genetic immunization experiments in animal models of infection with E. faecalis. Such fragments will be particularly useful for identifying protein epitopes able to provoke a prophylactic or therapeutic immune response. This approach can allow for the subsequent preparation of monoclonal antibodies of particular value from the requisite organ of the animal successfully resisting or clearing infection for the development of prophylactic agents or therapeutic treatments of enterococcal infection in mammals, particularly humans.

The hyperimmune serum reactive antigens and fragments thereof may be used as an antigen for vaccination of a host to produce specific antibodies which protect against invasion of bacteria, for example by blocking adherence of bacteria to damaged tissue. Examples of tissue damage include wounds in skin or connective tissue caused e.g. by mechanical, chemical or thermal damage or by implantation of indwelling devices, or wounds in the mucous membranes, such as the mouth, mammary glands, urethra or vagina.

The present invention also includes a vaccine formulation which comprises the immunogenic recombinant protein together with a suitable carrier. Since the protein may be broken down in the stomach, it is preferably administered parenterally, including, for example, administration that is subcutaneous, intramuscular, intravenous, intradermal intranasal or transdermal. Formulations suitable for parenteral administration include aqueous and non-aqueous sterile injection solutions which may contain anti-oxidants, buffers, bacteriostats and solutes which render the formulation isotonic with the bodily fluid, preferably the blood, of the individual; and aqueous and non-aqueous sterile suspensions which may include suspending agents or thickening agents. The formulations may be presented in unit-dose or multi-dose containers, for example, sealed ampoules and vials, and may be stored in a freeze-dried condition requiring only the addition of the sterile liquid carrier immediately prior to use. The vaccine formulation may also include adjuvant systems for enhancing the immunogenicity of the formulation, such as oil-in-water systems and other systems known in the art. The dosage will depend on the specific activity of the vaccine and can be readily determined by routine experimentation.

According to another aspect, the present invention relates to a pharmaceutical composition comprising such a hyperimmune serum-reactive antigen or a fragment thereof as provided in the present invention for E. faecalis. Such a pharmaceutical composition may comprise one, preferably at least two, or more hyperimmune serum reactive antigens or fragments thereof against E. faecalis. Optionally, such E. faecalis hyperimmune serum reactive antigens or fragments thereof may also be combined with antigens against other pathogens in a combination pharmaceutical composition.   Preferably, said pharmaceutical composition is a vaccine for preventing or treating an infection caused by E. faecalis and/or other pathogens against which the antigens have been included in the vaccine.

According to a further aspect, the present invention relates to a pharmaceutical composition comprising a nucleic acid molecule encoding a hyperimmune serum-reactive antigen or a fragment thereof as identified above for E. faecalis. Such a pharmaceutical composition may comprise one or more nucleic acid molecules encoding hyperimmune serum reactive antigens or fragments thereof against E. faecalis. Optionally, such E. faecalis nucleic acid molecules encoding hyperimmune serum reactive antigens or fragments thereof may also be combined with nucleic acid molecules encoding antigens against other pathogens in a combination pharmaceutical composition. Preferably, said pharmaceutical composition is a vaccine for preventing or treating an infection caused by E. faecalis and/or other pathogens against which the antigens have been included in the vaccine.

The pharmaceutical composition may contain any suitable auxiliary substances, such as buffer substances, stabilisers or further active ingredients, especially ingredients known in connection of pharmaceutical composition and/or vaccine production.

A preferable carrier/or excipient for the hyperimmune serum-reactive antigens, fragments thereof or a coding nucleic acid molecule thereof according to the present invention is an immunostimulatory compound for further stimulating the immune response to the given hyperimmune serum-reactive antigen, fragment thereof or a coding nucleic acid molecule thereof. Preferably the immunostimulatory compound in the pharmaceutical preparation according to the present invention is selected from the group of polycationic substances, especially polycationic peptides, immunostimulatory nucleic acids molecules, preferably immunostimulatory deoxynucleotides, alum, Freund's complete adjuvants, Freund's incomplete adjuvants, neuroactive compounds, especially human growth hormone, or combinations thereof.

It is also within the scope of the present invention that the pharmaceutical composition, especially vaccine, comprises apart from the hyperimmune serum reactive antigens, fragments thereof and/or coding nucleic acid molecules thereof according to the present invention other compounds which are biologically or pharmaceutically active. Preferably, the vaccine composition comprises at least one polycationic peptide. The polycationic compound(s) to be used according to the present invention may be any polycationic compound which shows the characteristic effects according to the WO 97/30721. Preferred polycationic compounds are selected from basic polypetides, organic polycations, basic polyamino acids or mixtures thereof. These polyamino acids should have a chain length of at least 4 amino acid residues (WO 97/30721). Especially preferred are substances like polylysine, polyarginine and polypeptides containing more than 20 %, especially more than 50 % of basic amino acids in a range of more than 8, especially more than 20, amino acid residues or mixtures thereof. Other preferred polycations and their pharmaceutical compositions are described in WO 97/30721 (e.g. polyethyleneimine) and WO 99/38528. Preferably these polypeptides contain between 20 and 500 amino acid residues, especially between 30 and 200 residues.

These polycationic compounds may be produced chemically or recombinantly or may be derived from natural sources.

Cationic (poly)peptides may also be anti-microbial with properties as reviewed in {Ganz, T., 1999}. These (poly)peptides may be of prokaryotic or animal or plant origin or may be produced chemically or recombinantly (WO 02/13857). Peptides may also belong to the class of defensins (WO 02/13857). Sequences of such peptides can be, for example, found in the Antimicrobial Sequences Database under the following internet address:

http://www.bbcm.univ.trieste.it/~tossi/pag2.html

Such host defence peptides or defensives are also a preferred form of the polycationic polymer according to the present invention. Generally, a compound allowing as an end product activation (or down-regulation) of the adaptive immune system, preferably mediated by APCs (including dendritic cells) is used as polycationic polymer.

Especially preferred for use as polycationic substances in the present invention are cathelicidin derived antimicrobial peptides or derivatives thereof (International patent application WO 02/13857, incorporated herein by reference), especially antimicrobial peptides derived from mammalian cathelicidin, preferably from human, bovine or mouse.

Polycationic compounds derived from natural sources include HIV-REV or HIV-TAT (derived cationic peptides, antennapedia peptides, chitosan or other derivatives of chitin) or other peptides derived from these peptides or proteins by biochemical or recombinant production. Other preferred polycationic compounds are cathelin or related or derived substances from cathelin. For example, mouse cathelin is a peptide which has the amino acid sequence NH2-RLAGLLRKGGEKIGEKLKKIGOKIKNFFQKLVPQPE-

COOH. Related or derived cathelin substances contain the whole or parts of the cathelin sequence with at least 15-20 amino acid residues. Derivations may include the substitution or modification of the natural amino acids by amino acids which are not among the 20 standard amino acids. Moreover, further cationic residues may be introduced into such cathelin molecules. These cathelin molecules are preferred to be combined with the antigen. These cathelin molecules surprisingly have turned out to be also effective as an adjuvant for an antigen without the addition of further adjuvants. It is therefore possible to use such cathelin molecules as efficient adjuvants in vaccine formulations with or without further immunactivating substances.

Another preferred polycationic substance to be used according to the present invention is a synthetic peptide containing at least 2 KLK-motifs separated by a linker of 3 to 7 hydrophobic amino acids (International patent application WO 02/32451, incorporated herein by reference).

The pharmaceutical composition of the present invention may further comprise immunostimulatory nucleic acid(s). Immunostimulatory nucleic acids are e. g. neutral or artificial CpG containing nucleic acids, short stretches of nucleic acids derived from non-vertebrates or in form of short oligonucleotides (ODNs) containing non-methylated cytosine-guanine di-nucleotides (CpG) in a certain base context (e.g. described in WO 96/02555). Alternatively, also nucleic acids based on inosine and cytidine as e.g. described in the WO 01/93903, or deoxynucleic acids containing deoxy-inosine and/or deoxyuridine residues (described in WO 01/93905 and PCT/EP 02/05448, incorporated herein by reference) may preferably be used as immunostimulatory nucleic acids for the present invention. Preferably, the mixtures of different immunostimulatory nucleic acids may be used according to the present invention.

It is also within the present invention that any of the aforementioned polycationic compounds is combined with any of the immunostimulatory nucleic acids as aforementioned. Preferably, such combinations are according to the ones as described in WO 01/93905, WO 02/32451, WO 01/54720, WO 01/93903, WO 02/13857 and PCT/EP 02/05448 and the Austrian patent application A 1924/2001, incorporated herein by reference.

In addition or alternatively such vaccine composition may comprise apart from the hyperimmune serum reactive antigens and fragments thereof, and the coding nucleic acid molecules thereof according to the present invention a neuroactive compound. Preferably, the neuroactive compound is human growth factor as, e.g. described in WO 01/24822. Also preferably, the neuroactive compound is combined with any of the polycationic compounds and/or immunostimulatory nucleic acids as afore-mentioned.

In a further aspect the present invention is related to a pharmaceutical composition. Such pharmaceutical composition is, for example, the vaccine described herein. Also a pharmaceutical composition is a pharmaceutical composition which comprises any of the following compounds or combinations thereof: the nucleic acid molecules according to the present invention, the hyperimmune serum reactive antigens and fragments thereof according to the present invention, the vector according to the present invention, the cells according to the present invention, the antibody according to the present invention, the functional nucleic acids according to the present invention and the binding peptides such as the anticalines according to the present invention, any agonists and antagonists screened as described herein. In connection therewith any of these compounds may be employed in combination with a non-sterile or sterile carrier or carriers for use with cells, tissues or organisms, such as a pharmaceutical carrier suitable for administration to a subject. Such compositions comprise, for instance, a media additive or a therapeutically effective amount of a hyperimmune serum reactive antigen and fragments thereof of the invention and a pharmaceutically acceptable carrier or excipient. Such carriers may include, but are not limited to, saline, buffered saline, dextrose, water, glycerol, ethanol and combinations thereof. The formulation should suit the mode of administration.

The pharmaceutical compositions may be administered in any effective, convenient manner including, for instance, administration by topical, oral, anal, vaginal, intravenous, intraperitoneal, intramuscular, subcutaneous, intranasal or intradermal routes among others.

In therapy or as a prophylactic, the active agent may be administered to an individual as an injectable composition, for example as a sterile aqueous dispersion, preferably isotonic.

Alternatively the composition may be formulated for topical application, for example in the form of ointments, creams, lotions, eye ointments, eye drops, ear drops, mouthwash, impregnated dressings and sutures and aerosols, and may contain appropriate conventional additives, including, for example, preservatives, solvents to assist drug penetration, and emollients in ointments and creams. Such topical formulations may also contain compatible conventional carriers, for example cream or ointment bases, and ethanol or oleyl alcohol for lotions. Such carriers may constitute from about 1 % to about 98 % by weight of the formulation; more usually they will constitute up to about 80 % by weight of the formulation.

In addition to the therapy described above, the compositions of this invention may be used generally as a wound treatment agent to prevent adhesion of bacteria to matrix proteins exposed in wound tissue and for prophylactic use in dental treatment as an alternative to, or in conjunction with, antibiotic prophylaxis.

A vaccine composition is conveniently in injectable form. Conventional adjuvants may be employed to enhance the immune response. A suitable unit dose for vaccination is 0.05-5 µg antigen / per kg of body weight, and such dose is preferably administered 1-3 times and with an interval of 1-3 weeks.

With the indicated dose range, no adverse toxicological effects should be observed with the compounds of the invention which would preclude their administration to suitable individuals.

In a further embodiment the present invention relates to diagnostic and pharmaceutical packs and kits comprising one or more containers filled with one or more of the ingredients of the aforementioned compositions of the invention. The ingredient(s) can be present in a useful amount, dosage, formulation or combination. Associated with such container(s) can be a notice in the form prescribed by a governmental agency regulating the manufacture, use or sale of pharmaceuticals or biological products, reflecting approval by the agency of the manufacture, use or sale of the product for human administration.

In connection with the present invention any disease related use as disclosed herein such as, e. g. use of the pharmaceutical composition or vaccine, is particularly a disease or diseased condition which is caused by, linked or associated with Enterococci, more preferably, E. faecalis. In connection therewith it is to be noted that E. faecalis comprises several strains including those disclosed herein. A disease related, caused or associated with the bacterial infection to be prevented and/or treated according to the present invention includes besides others bacterial pharyngitis, scarlet fever, impetigo, rheumatic fever, necrotizing fasciitis and sepsis in humans.

In a still further embodiment the present invention is related to a screening method using any of the hyperimmune serum reactive antigens or nucleic acids according to the present invention. Screening methods as such are known to the one skilled in the art and can be designed such that an agonist or an antagonist is screened. Preferably an antagonist is screened which in the present case inhibits or prevents the binding of any hyperimmune serum reactive antigen and fragment thereof according to the present invention to an interaction partner. Such interaction partner can be a naturally occurring interaction partner or a non-naturally occurring interaction partner.

The invention also provides a method of screening compounds to identify those which enhance (agonist) or block (antagonist) the function of hyperimmune serum reactive antigens and fragments thereof or nucleic acid molecules of the present invention, such as its interaction with a binding molecule. The method of screening may involve high-throughput.

For example, to screen for agonists or antagonists, the interaction partner of the nucleic acid molecule and nucleic acid, respectively, according to the present invention, maybe a synthetic reaction mix, a cellular compartment, such as a membrane, cell envelope or cell wall, or a preparation of any thereof, may be prepared from a cell that expresses a molecule that binds to the hyperimmune serum reactive antigens and fragments thereof of the present invention. The preparation is incubated with labelled hyperimmune serum reactive antigens and fragments thereof in the absence or the presence of a candidate molecule which may be an agonist or antagonist. The ability of the candidate molecule to bind the binding molecule is reflected in decreased binding of the labelled ligand. Molecules which bind gratuitously, i. e., without inducing the functional effects of the hyperimmune serum reactive antigens and fragments thereof, are most likely to be good antagonists. Molecules that bind well and elicit functional effects that are the same as or closely related to the hyperimmune serum reactive antigens and fragments thereof are good agonists.

The functional effects of potential agonists and antagonists may be measured, for instance, by determining the activity of a reporter system following interaction of the candidate molecule with a cell or appropriate cell preparation, and comparing the effect with that of the hyperimmune serum reactive antigens and fragments thereof of the present invention or molecules that elicit the same effects as the hyperimmune serum reactive antigens and fragments thereof. Reporter systems that may be useful in this regard include but are not limited to colorimetric labelled substrate converted into product, a reporter gene that is responsive to changes in the functional activity of the hyperimmune serum reactive antigens and fragments thereof, and binding assays known in the art.

Another example of an assay for antagonists is a competitive assay that combines the hyperimmune serum reactive antigens and fragments thereof of the present invention and a potential antagonist with membrane-bound binding molecules, recombinant binding molecules, natural substrates or ligands, or substrate or ligand mimetics, under appropriate conditions for a competitive inhibition assay. The hyperimmune serum reactive antigens and fragments thereof can be labelled such as by radioactivity or a colorimetric compound, such that the molecule number of hyperimmune serum reactive antigens and fragments thereof bound to a binding molecule or converted to product can be determined accurately to assess the effectiveness of the potential antagonist.

Potential antagonists include small organic molecules, peptides, polypeptides and antibodies that bind to a hyperimmune serum reactive antigen and fragments thereof of the invention and thereby inhibit or extinguish its acitivity. Potential antagonists also may be small organic molecules, a peptide, a polypeptide such as a closely related protein or antibody that binds to the same sites on a binding molecule without inducing functional activity of the hyperimmune serum reactive antigens and fragments thereof of the invention.

Potential antagonists include a small molecule which binds to and occupies the binding site of the hyperimmune serum reactive antigens and fragments thereof thereby preventing binding to cellular binding molecules, such that normal biological activity is prevented. Examples of small molecules include but are not limited to small organic molecules, peptides or peptide-like molecules.

Other potential antagonists include antisense molecules (see {Okano, H. et al., 1991}; OLIGODEOXYNUCLEOTIDES AS ANTISENSE INHIBITORS OF GENE EXPRESSION; CRC Press, Boca Ration, FL (1988), for a description of these molecules).

Preferred potential antagonists include derivatives of the hyperimmune serum reactive antigens and fragments thereof of the invention.

As used herein the activity of a hyperimmune serum reactive antigen and fragment thereof according to the present invention is its capability to bind to any of its interaction partner or the extent of such capability to bind to its or any interaction partner.

In a particular aspect, the invention provides the use of the hyperimmune serum reactive antigens and fragments thereof, nucleic acid molecules or inhibitors of the invention to interfere with the initial physical interaction between a pathogen and mammalian host responsible for sequelae of infection. In particular the molecules of the invention may be used: i) in the prevention of adhesion of E. faecalis to mammalian extracellular matrix proteins on in-dwelling devices or to extracellular matrix proteins in wounds; ii) to block protein mediated mammalian cell invasion by, for example, initiating phosphorylation of mammalian tyrosine kinases (Rosenshine, I. et al., 1992) to block bacterial adhesion between mammalian extracellular matrix proteins and bacterial proteins which mediate tissue damage; iv) to block the normal progression of pathogenesis in infections initiated other than by the implantation of in-dwelling devices or by other surgical techniques.

Each of the DNA coding sequences provided herein may be used in the discovery and development of antibacterial compounds. The encoded protein upon expression can be used as a target for the screening of antibacterial drugs. Additionally, the DNA sequences encoding the amino terminal regions of the encoded protein or Shine-Delgarno or other translation facilitating sequences of the respective mRNA can be used to construct antisense sequences to control the expression of the coding sequence of interest.

The antagonists and agonists may be employed, for instance, to inhibit diseases arising from infection with Enterococci, especially E. faecalis, such as sepsis.

In a still further aspect the present invention is related to an affinity device such affinity device comprises as least a support material and any of the hyperimmune serum reactive antigens and fragments thereof according to the present invention, which is attached to the support material. Because of the specificity of the hyperimmune serum reactive antigens and fragments thereof according to the present invention for their target cells or target molecules or their interaction partners, the hyperimmune serum reactive antigens and fragments thereof allow a selective removal of their interaction partner(s) from any kind of sample applied to the support material provided that the conditions for binding are met. The sample may be a biological or medical sample, including but not limited to, fermentation broth, cell debris, cell preparation, tissue preparation, organ preparation, blood, urine, lymph liquid, liquor and the like.

The hyperimmune serum reactive antigens and fragments thereof may be attached to the matrix in a covalent or non-covalent manner. Suitable support material is known to the one skilled in the art and can be selected from the group comprising cellulose, silicon, glass, aluminium, paramagnetic beads, starch and dextrane.

The present invention is further illustrated by the following figures, examples and the sequence listing from which further features, embodiments and advantages may be taken. It is to be understood that the present examples are given by way of illustration only and not by way of limitation of the disclosure.

In connection with the present invention

Figure 1 shows the characterization of E. faecalis specific human sera.

Figure 2 shows the characterization of the small fragment genomic library, LEF-70, from Enterococcus faecalis V583.

Figure 3 shows the selection of bacterial cells by MACS using biotinylated human IgGs.

Table 1 shows the summary of all screens performed with genomic *E. faecalis* libraries and human serum and *E. faecium* proteins identified by homology search.

Table 2 shows the epitope serology with human sera.

The figures to which it might be referred to in the specification are described in the following in more details.

Figure 1 shows the characterization of human sera for anti-*E. faecalis* antibodies as measured by immune assays. Total anti- *E. faecalis* IgG antibody levels were measured by standard ELISA using total bacterial lysates or culture supernant fractions prepared from *E. faecalis* strain V583 as coating antigens. (A) Results of representative experiments are shown with healthy adult sera. Data are expressed as ELISA units calculated from absorbance at 405nm at two serum dilutions in the linear range of detection (2.000X and 10,000X). (B) Immunoblot analysis was performed with high titer sera from healthy adults selected by ELISA in order to ensure multiple immune reactivity with protein antigens. Results of a representative experiment using total bacterial lysate and selected human sera at 5.000X dilution are shown. Blots were developed with anti-human IgG secondary antibody reagent. Lanes 1-4: individual high titer sera included in screening pools (N1-4), lane 5: low titer serum (N5), lane 6: negative control, (no serum, 2nd antibody only). Mw: molecular weight markers. (C) shows the comparison of IgG titers obtained with sera from healthy adults (N, numbering is the same as for (B)) vs. acute phase patients (P) with enterococcal infections (such as wound infection, bacteremia, infected catheter). Data are expressed as ELISA units, and were calculated as for (A). (D) shows the results of experiments with convalescent sera from endocarditis patients. Data are expressed as ELISA units calculated from absorbance at 405nm at serum dilution in the linear range of detection (2.000X). (E) Immunoblot analysis was performed with high titer sera from endocarditis patients in order to ensure multiple immune reactivity with protein antigens. Results of a representative experiment using total bacterial lysate and selected human sera at 5.000X dilution are shown. Blots were developed with anti-human IgG secondary antibody reagent. Lanes 1-16: individual sera from endocarditis patients. Mw: molecular weight markers.

Figure 2 (A) shows the fragment size distribution of the *Enterococcus faecalis* V583 small fragment genomic library, LEF-70. After sequencing 576 randomly selected clones, sequences were trimmed to eliminate vector residues and the numbers of clones with various genomic fragment sizes were plotted. (B) shows the graphic illustration of the distribution of the set of 483 randomly sequenced clones of LEF-70 over the *Enterococcus faecalis* V583 chromosome. Rectangles indicate matching sequences to annotated ORFs and diamonds represent fully matched clones to non-coding chromosomal sequences in +/+ or +/- orientation, respectively. Circles position all clones with chimeric sequences. Numeric distances in base pairs are indicated over the circular genome for orientation. Partitioning of various clone sets within the library is given in numbers and percentage at the bottom of the figure.

Figure 3A shows the MACS selection with biotinylated human IgGs. The LEF-70 library in pMAL9.1 was screened with 20 µg biotinylated, human serum (IC10-IgG) in the first and in the second selection round. As negative control, no serum was added to the library cells for screening. Number of cells selected after the 1st and 2nd elution for round 1 and 1st to 4th elution for round 2 are shown for the selection with and without added IgGs. Figure 3B shows the reactivity of specific clones (1-26) isolated after 2 rounds of bacterial surface display as analysed by Western blot analysis with the human serum (IC10-IgG, approximately 4 µg/µl) used for selection by MACS at a dilution of 1:3,000. As a loading control the same blot was also analysed with antibodies directed against the platform protein LamB at a dilution of 1:5,000. LB, Extract from a clone expressing LamB without foreign peptide insert.

Table 1: Immunogenic proteins identified by bacterial surface display.
1a shows the antigens identified with sera from healthy adults with high anti-enterococcus titer. A, LEF-300 library from *Enterococcus faecalis* V583 in *fhuA* with IC10-IgG (723), B, LEF-300 library in *fhuA* with IC9-IgG (389), C, LEF-70 library in *lamB* with IC10-IgG (1096), D, LEF-70 library in *lamB* with IC9-IgG (1065). Table 1b shows *E. faecium* proteins identified by homology search with a minimum of 70% identity to *E. faecalis* antigens listed in table 1a. The computer program TBLASTN was used to determine identity between   E.   *faecalis*   and   E.   *faecium*   sequences (http://www.hgsc.bcm.tmc.edu/microbial/microbialblast.cgi?organism=Efaecium). The amino acid sequence identity was calculated based on the complete ORF sequence of the corresponding *E. faecalis* antigen. 1c shows the antigens identified with sera from convalescing endocarditis patients with high anti-enterococcus titer. E, LEF-70 library from *Enterococcus faecalis* V583 in LamB with P25-IgG (843); F, LEF-70 library in LamB with P26-IgG (845); G, LEF-300 library in FhuA with P25-IgG (691); H, LEF-300 library in FhuA with P26-IgG (770); *, prediction of antigenic sequences longer than 5 amino acids was performed with the program ANTIGENIC {Kolaskar, A. et al., 1990}.

Table 2: Epitope serology with human sera.
Immune reactivity of individual synthetic peptides representing selected epitopes with individual human sera is shown. Extent of reactivity is colour coded; white: - (<100 U), light grey: + (100-249 U), dark grey: ++ (250-349 U), black: +++ (>350 U). ELISA units (U) are calculated from OD_{405nm} readings and the serum dilution. Score is calculated as the sum of all reactivities (- = 0; + = 1; ++ = 2; +++ = 3). N1 to N10 sera are high titer sera from healthy adults used in the screens with IC9- and IC10 IgG pools. P1-P11 are sera from patients with *E. faecalis* infections. Location of synthetic peptides within the antigenic ORFs according to the genome annotation of V583 strain are given indicating the first and last amino acid residues. Peptide names: EF0020.1 present in annotated ORF EF0020 on the chromosome; ARF0679.1, potential novel ORF in alternative reading-frame of EF0679, EFC0034.1 present in annotated ORF from plasmid C; ARFC0021.1 present in potential novel ORF in alternative reading-frame of ARFC0021 from plasmid C.

EXAMPLES

Example 1: Characterization and selection of human serum sources based on anti-*E. faecalis* antibodies, preparation of antibody screening reagents

Experimental procedures

Enzyme linked immune assay (ELISA).
ELISA plates (Maxisorb, Millipore) were coated with 5-10 µg/ml total protein diluted in coating buffer (0.1M sodium carbonate pH 9.2). Three dilutions of sera (2,000X, 10,000X, 50,000X) were made in PBS-BSA. Highly specific Horse Radish Peroxidase (HRP)-conjugated anti-human IgG or anti-human IgA secondary antibodies (Southern Biotech) were used according to the manufacturers' recommendations (dilution: 1,000x). Antigen-antibody complexes were quantified by measuring the conversion of the substrate (ABTS) to colored product based on OD_{405nm} readings by automatic ELISA reader (TECAN SUNRISE).

Preparation of bacterial antigen extracts
Total bacterial lysate: Bacteria were grown overnight in BHI (Brain-heart Infusion) and lysed by repeated freeze-thaw cycles: incubation on dry ice/ethanol-mixture until frozen (1 min), then thawed at 37°C (5 min): repeated 3 times. This was followed by sonication and collection of supernatant by centrifugation (3,500 rpm, 15 min, 4°C).

Culture supernatant: After removal of bacteria by centrifugation, the supernatant of overnight grown bacterial cultures was precipitated with ice-cold ethanol by mixing 1 part supernatant with 3 parts absolute ethanol and incubated overnight at -20°C. Precipitates were collected by centrifugation (2,600 g, for 15 min). Dry pellets were dissolved either in PBS for ELISA, or in urea and SDS-sample buffer for SDS-PAGE and immunoblotting. The protein concentration of samples was determined by Bradford assay.

Immunoblotting

Total bacterial lysate and culture supernatant samples were prepared from *in vitro* grown *E. faecalis* strain V583. 10 to 25μg total protein/lane was separated by SDS-PAGE using the BioRad Mini-Protean Cell electrophoresis system and proteins transferred to nitrocellulose membrane (ECL, Amersham Pharmacia). After overnight blocking in 5% milk, human sera were added at 2,000x dilution, and HRPO labeled anti-human IgG was used for detection.

Purification of antibodies for genomic screening. Five sera from both the patient and the healthy group were selected based on the overall anti-*E. faecalis* titers for serum pools used in the screening procedure. Antibodies against *E. coli* proteins were removed by incubating the heat-inactivated sera with whole cell *E. coli* cells (DH5alpha, transformed with pHIE11, grown under the same condition as used for bacterial surface display). Highly enriched preparations of IgGs from the pooled, depleted sera were generated by protein G affinity chromatography, according to the manufacturer's instructions (UltraLink Immobilized Protein G, Pierce). The efficiency of depletion and purification was checked by SDS-PAGE, Western blotting, ELISA and protein concentration measurements.

Results

The antibodies produced against *E. faecalis* by the human immune system and present in human sera are indicative of the *in vivo* expression of the antigenic proteins and their immunogenicity. These molecules are essential for the identification of individual antigens in the approach as described in the present invention, which is based on the interaction of the specific anti-enterococcal antibodies and the corresponding Enterococcal peptides or proteins. To gain access to relevant antibody repertoires, human sera were collected from healthy adults, as well from patients with enterococcal infections and naïve individuals, young children between 5 and 10 months of age, after they already lost maternal antibodies.

Antibodies in serum and other body fluids, induced in individuals exposed to the pathogens are crucial for antigen identification. The exposure to Enterococci results in asymptomatic colonization, current or past acute or chronic infection. *E. faecalis* colonization and infections are common, and antibodies are present as a consequence of natural immunization from previous encounters. Since Enterococci are opportunistic (not obliguate) pathogens, it is likely that sera from healthy individuals also contain relevant antibodies.

70 sera from healthy adults, 16 from patients convalescing from endocarditis caused by *E. faecalis* and 8 from patients with acute infection (mainly wound and bacteremia) were characterized for anti-*E. faecalis* antibodies by a series of immune assays. Primary characterization was done by ELISA using two different antigen preparations, such as total bacterial extract and culture supernatant proteins prepared from *E. faecalis* V583 strain. Representative experiments are shown in Fig. 1 using sera from the healthy adult population (Fig. 1A) and from endocarditis patients (Fig.1D). Antibody titers were compared at given dilutions where the response was linear. Sera were ranked based on the IgG reactivity against the two complex antigenic mixtures, and the highest ones were selected for further testing by immunoblotting. This analysis confirmed high antibody reactivity of the pre-selected sera against multiple Enterococcal proteins (Fig 1B&E), especially when compared to not selected, low-titer sera. ELISA ranking of sera did not always correlated with immunoblot signals suggesting that antibodies against non-protein components (e.g. lipoteichoic acid, peptidoglycan, etc.) contributed to the total ELISA

reactivities against total bacterial extracts. Thus the final selection of sera to be included in antibody-pools was based mainly on multiple immunogenic bands in immunoblotting experiments. This extensive antibody characterization approach has led to the unambiguous identification of anti-*E. faecalis* hyperimmune sera. These selected sera from healthy adults were compared to sera from patients with various acute enterococcal infections, such as wound infection, bacteremia and catheter related infections. It was obvious that anti-*E. faecalis* antibody titer was lower in the acute patient group, suggesting that disease might occur in low titer individuals (Fig.1C). These patient sera were then used mainly for validation purposes, that is measuring antibody levels against identified epitopes. However, patients with endocarditis caused by *E. faecalis* (verified by routine microbiological diagnosis) developed high levels of antibodies in the convalescent phase, as it is demonstrated by the high percentage of high titer sera among the 16 samples analysed here (Fig. 1E&D).

10 sera were selected from the healthy adults and 10 from the convalescent patients with endocarditis donor groups. Sera were pooled (5 samples/pool), and IgG purified for antigen identification by bacterial surface display IgG antibodies were purified from pooled sera by affinity chromatography and depleted of *E. coli* -reactive antibodies to avoid background in the bacterial surface display screen (two IgG pools: NEf9, NEf10).

Example 2: Generation of highly random, frame-selected, small-fragment, genomic DNA libraries of *Enterococcus faecalis*

Experimental procedures

Preparation of enterococcal genomic DNA. 50 ml Brain heart infusion (BHI) medium was inoculated with *Enterococcus faecalis* V583 bacteria from a frozen stab and grown with aeration and shaking for 18 h at 37°C. The culture was then harvested, centrifuged with 1,600x g for 15 min and the supernatant was removed. Bacterial pellets were washed 3 x with PBS and carefully re-suspended in 0.5 ml of Lysozyme solution (100 mg/ml). 0.1 ml of 10 mg/ml heat treated RNase A and 20 U of RNase T1 were added, mixed carefully and the solution was incubated for 1 h at 37°C. Following the addition of 0.2 ml of 20 % SDS solution and 0.1 ml of Proteinase K (10 mg/ml) the tube was incubated overnight at 55°C. 1/3 volume of saturated NaCl was then added and the solution was incubated for 20 min at 4°C. The extract was pelleted in a microfuge (13,000 rpm) and the supernatant transferred into a new tube. The solution was extracted with PhOH/CHCl$_3$/IAA (25:24:1) and with CHCl$_3$/IAA (24:1). DNA was precipitated at room temperature by adding 0.6x volume of Isopropanol, spooled from the solution with a sterile Pasteur pipette and transferred into tubes containing 80% ice-cold ethanol. DNA was recovered by centrifuging the precipitates with 10-12,000x g, then dried on air and dissolved in ddH$_2$O.

Preparation of small genomic DNA fragments. Genomic DNA fragments were mechanically sheared into fragments ranging in size between 150 and 300 bp using a cup-horn sonicator (Bandelin Sonoplus UV 2200 sonicator equipped with a BB5 cup horn, 10 sec. pulses at 100 % power output) or into fragments of size between 50 and 70 bp by mild DNase I treatment (Novagen). It was observed that sonication yielded a much tighter fragment size distribution when breaking the DNA into fragments of the 150-300 bp size range. However, despite extensive exposure of the DNA to ultrasonic wave-induced hydromechanical shearing force, subsequent decrease in fragment size could not be efficiently and reproducibly achieved. Therefore, fragments of 50 to 70 bp in size were obtained by mild DNase I treatment using Novagen's shotgun cleavage kit. A 1:20 dilution of DNase I provided with the kit was prepared and the digestion was performed in the presence of MnCl$_2$ in a 60 µl volume at 20°C for 5 min to ensure double-stranded cleavage by the enzyme. Reactions were stopped with 2 µl of 0.5 M EDTA and the fragmentation efficiency was evaluated on a 2% TAE-agarose gel. This treatment resulted in total fragmentation of genomic DNA into near 50-70 bp fragments. Fragments were then blunt-ended twice using T4 DNA Polymerase in the presence of 100 µM each of dNTPs to ensure efficient flushing of the ends. Fragments

were used immediately in ligation reactions or frozen at -20°C for subsequent use.

Description of the vectors. The vector pMAL4.31 was constructed on a pASK-IBA backbone {Skerra, A., 1994} with the beta-lactamase (*bla*) gene exchanged with the Kanamycin resistance gene. In addition the *bla* gene was cloned into the multiple cloning site. The sequence encoding mature beta-lactamase is preceded by the leader peptide sequence of *ompA* to allow efficient secretion across the cytoplasmic membrane. Furthermore a sequence encoding the first 12 amino acids (spacer sequence) of mature beta-lactamase follows the *ompA* leader peptide sequence to avoid fusion of sequences immediately after the leader peptidase cleavage site, since e.g. clusters of positive charged amino acids in this region would decrease or abolish translocation across the cytoplasmic membrane {Kajava, A. et al., 2000}. A *SmaI* restriction site serves for library insertion. An upstream FseI site and a downstream NotI site, which were used for recovery of the selected fragment, flank the *SmaI* site. The three restriction sites are inserted after the sequence encoding the 12 amino acid spacer sequence in such a way that the *bla* gene is transcribed in the -1 reading frame resulting in a stop codon 15 bp after the NotI site. A +1 bp insertion restores the *bla* ORF so that beta-lactamase protein is produced with a consequent gain of Ampicillin resistance.

The vector pMAL9.1 was constructed by cloning the *lamB* gene into the multiple cloning site of pEH1 {Hashemzadeh-Bonehi, L. et al., 1998}. Subsequently, a sequence was inserted in *lamB* after amino acid 154, containing the restriction sites FseI, SmaI and NotI. The reading frame for this insertion was constructed in such a way that transfer of frame-selected DNA fragments excised by digestion with FseI and NotI from plasmid pMAL4.31 yields a continuous reading frame of *lamB* and the respective insert.

The vector pHIE11 was constructed by cloning the *fhuA* gene into the multiple cloning site of pEH1. Thereafter, a sequence was inserted in *fhuA* after amino acid 405, containing the restriction site FseI, XbaI and NotI. The reading frame for this insertion was chosen in a way that transfer of frame-selected DNA fragments excised by digestion with FseI and NotI from plasmid pMAL4.31 yields a continuous reading frame of *fhuA* and the respective insert.

Cloning and evaluation of the library for frame selection. Genomic *E. faecalis* DNA fragments were ligated into the *SmaI* site of the vector pMAL4.31. Recombinant DNA was electroporated into DH10B electrocompetent *E. coli* cells (GIBCO BRL) and transformants plated on LB-agar supplemented with Kanamycin (50 µg/ml) and Ampicillin (50 µg/ml). Plates were incubated over night at 37°C and colonies collected for large scale DNA extraction. A representative plate was stored and saved for collecting colonies for colony PCR analysis and large-scale sequencing. A simple colony PCR assay was used to initially determine the rough fragment size distribution as well as insertion efficiency. From sequencing data the precise fragment size was evaluated, junction intactness at the insertion site as well as the frame selection accuracy (3*n*+1 rule).

Cloning and evaluation of the library for bacterial surface display. Genomic DNA fragments were excised from the pMAL4.31 vector, containing the *E. faecalis* library with the restriction enzymes FseI and NotI. The entire population of fragments was then transferred into plasmids pMAL9.1 (LamB) or pHIE11 (FhuA), which have been digested with FseI and NotI. Using these two restriction enzymes, which recognise an 8 bp GC rich sequence, the reading frame that was selected in the pMAL4.31 vector is maintained in each of the platform vectors. The plasmid library was then transformed into *E. coli* DH5alpha cells by electroporation. Cells were plated onto large LB-agar plates supplemented with 50 µg/ml Kanamycin and grown over night at 37°C at a density yielding clearly visible single colonies. Cells were then scraped off the surface of these plates, washed with fresh LB medium and stored in aliquots for library screening at -80°C.

Results

Libraries for frame selection. Three libraries (LEF-70 and LEF-300) were generated in the pMAL4.31 vector

with sizes of approximately 70 and 300 bp, respectively. For both library, ligation and subsequent transformation of approximately 1 μg of pMAL4.31 plasmid DNA and 50 ng of fragmented genomic *E. faecalis* DNA yielded approximately 1.5x 10^5 to 1x 10^6 clones after frame selection. To assess the randomness of the libraries, 576 randomly chosen clones of LEF-70 were sequenced. Of these sequences 483 were successfully trimmed and subjected to further bioinformatic analysis. The results showed that of these clones only very few were present more than once. Furthermore, it was shown that 73% of the clones fell in the size range between 25 and 100 bp and on average the clones had a size of 82 bp (Figure 2). Allmost all sequences followed the 3n+1 rule, showing that all clones were properly frame selected.

Bacterial surface display libraries. The display of peptides on the surface of *E. coli* required the transfer of the inserts from the LSPy libraries from the frame selection vector pMAL4.31 to the display plasmids pMAL9.1 (LamB) or pHIE11 (FhuA). Genomic DNA fragments were excised by *FseI* and *NotI* restriction and ligation of 5 ng inserts with 0.1 μg plasmid DNA and subsequent transformation into DH5alpha cells resulted in 7x 10^5 to 2x 10^6 clones. The clones were scraped off the LB plates and frozen without further amplification. These cells served as libraries for the subsequent screening procedure.

Example 3: Identification of highly immunogenic peptide sequences from *E. faecalis* using bacterial surface displayed genomic libraries and human serum

Experimental procedures

MACS screening. Approximately 2.5x 10^8 cells from a given library were grown in 5 ml LB-medium supplemented with 50 μg/ml Kanamycin for 2 h at 37°C. Expression was induced by the addition of 1 mM IPTG for 30 min. Cells were washed twice with fresh LB medium and approximately 2x 10^7 cells re-suspended in 100 μl LB medium and transferred to an Eppendorf tube.

20 μg of biotinylated, human IgGs purified from serum was added to the cells and the suspension incubated over night at 4°C with gentle shaking. 900 μl of LB medium was added, the suspension mixed and subsequently centrifuged for 10 min at 6,000 rpm at 4°C. Antibody reacting positive clones were captured with biotinylated anti-human-IgG secondary antibodies. Cells were washed once with 1 ml LB and then re-suspended in 100 μl LB medium. 10 μl of MACS microbeads coupled to streptavidin (Miltenyi Biotech, Germany) were added and the incubation continued for 20 min at 4°C. Thereafter 900 μl of LB medium was added and the MACS microbead cell suspension was loaded onto the equilibrated MS column (Miltenyi Biotech, Germany) which was fixed to the magnet. (The MS columns were equilibrated by washing once with 1 ml 70% EtOH and twice with 2 ml LB medium.)

The column was then washed three times with 3 ml LB medium. After removal of the magnet, cells were eluted by washing with 2 ml LB medium. After washing the column with 3 ml LB medium, the 2 ml eluate was loaded a second time on the same column and the washing and elution process repeated. In some cases the washing and elution process was performed up to four times to increase the specificity of the selection procedure, resulting in a final eluate of 2 ml.

A second round of screening was performed as follows. The cells from the final eluate were collected by centrifugation and re-suspended in 1 ml LB medium supplemented with 50 μg/ml Kanamycin. The culture was incubated at 37°C for 90 min and then induced with 1 mM IPTG for 30 min. Cells were subsequently collected, washed once with 1 ml LB medium and suspended in 10 μl LB medium. For the second round of screening the same amount (20 μg) of human, biotinylated IgGs was added and the suspension incubated over night at 4°C with gentle shaking. All further steps were exactly the same as in the first selection round, except that the number of washing, loading and elution cycles was adapted to each individual screening round. Cells selected after two rounds of selection were plated onto LB-agar plates supplemented with 50 μg/ml Kanamycin and grown over night at 37°C.

Evaluation of selected clones by sequencing and Western blot analysis. Selected clones were grown over night at 37°C in 3 ml LB medium supplemented with 50 µg/ml Kanamycin to prepare plasmid DNA using standard procedures. Sequencing was performed at MWG (Germany) or in collaboration with TIGR (U.S.A.).

For Western blot analysis approximately 10 to 20 µg of total cellular protein was separated by 10% SDS-PAGE and blotted onto HybondC membrane (Amersham Pharmacia Biotech, England). The LamB or FhuA fusion proteins were detected using human serum as the primary antibody at a dilution of approximately 1:3,000 to 1:5,000 and anti-human IgG antibodies coupled to HRP at a dilution of 1:5,000 as secondary antibodies. Detection was performed using the ECL detection kit (Amersham Pharmacia Biotech, England). Alternatively, rabbit anti FhuA or mouse anti LamB antibodies were used as primary antibodies in combination with the respective secondary antibodies coupled to HRP for the detection of the fusion proteins.

Results

Screening of bacterial surface display libraries by magnetic activated cell sorting (MACS) using biotinylated Igs. The libraries LEF-70 in pMAL9.1 and LEF-300 in pHIE11 were screened with pools of biotinylated, human IgGs from sera from healthy individuals (Table 1a) or from sera from convalescent endocarditis patients (Table 1c) (see Example 1: *Preparation of antibodies from human serum*). The selection procedure was performed as described under Experimental procedures. Figure 3A shows a representative example of a screen with the LEF-70 library and IC10-IgGs. As can be seen from the colony count after the first selection cycle from MACS screening, the total number of cells recovered at the end is drastically reduced from 1x 10^7 cells to approximately 5x 10^4 cells, whereas the selection without antibodies added showed a reduction to about 6x10^3 cells (Figure 3A). For the second round, 2x 10^5 cells were loaded on the column and almost all (1x 10^5) recovered with IC10-IgG after 4 cycles of loading, washing and elution, while fewer than 100 cells were recovered when no IgGs from human serum were added. This result clearly showed that selection was dependent on *E. faecalis* specific antibodies. Typically, the selection resulted in approximately 30 to 70% of clones to be selected specifically with the applied serum IgGs. For each screen between 400 and 1200 clones were subjected to DNA sequencing and the obtained sequence blasted against the genome of *Entereococcus faecalis* V583. This bioinformatic analysis allowed the evaluation of the screen on the basis of the frequency with which individual clones were selected by the addition of human IgGs. In order to confirm the good performance of the screen, all clones selected more than once were picked and subjected to Western blot analysis with the same, pooled serum IgGs (Figure 3B). The analysis of the more frequently selected clones (> 5x) revealed that more than 90% of these selected clones showed reactivity with antibodies present in the relevant serum whereas the control strain expressing LamB without a *E. faecalis* specific insert did not react with the same serum IgGs. In general, the clones which were selected with a lower frequency (< 5x), showed a lower percentage of reactivity with the applied serum IgGs. Colony PCR analysis showed that all selected and tested clones contained an insert in the expected size range.

The sequencing of a large number of randomly picked clones (400 to 1200 per screen) led also to the identification of the gene and the corresponding peptide or protein sequence that was specifically recognized by the human serum IgGs used for screening. The frequency with which a specific clone is selected reflects at least in part the abundance and/or affinity of the specific antibodies in the serum used for selection and recognizing the epitope presented by this clone. In that regard it is striking that clones derived from some ORFs (e.g. EF3060, EFA0042) were picked multiple times and in more than one screen, indicating their highly immunogenic property. Table 1a summarizes the data obtained with four screens performed with sera from healthy adults and Table 1c with four screens with convalescent endocarditis patients sera. The latter table contains only the antigens that were novel relative to the ones listed in Table 1a. 26 antigens were identified by both sera, 71 and 42 unique antigenic ORFs were selected by the sera from healthy adults and patients, respectively.

All clones that are presented in Table 1 have been verified by Western blot analysis using whole cellular extracts from single clones to show the indicated reactivity with the pool of human serum used in the respective screen. As can be seen from Table 1, distinct regions of the identified ORF are identified as immunogenic, since variably sized fragments of the proteins are displayed on the surface by the platform proteins.

It is further worth noticing that many of the genes identified by the bacterial surface display screen encode proteins that are either attached to the surface of E. faecalis and/or are secreted. This is in accordance with the expected role of surface attached or secreted proteins in virulence of E. faecalis.

Table 1b shows E. faecium proteins identified by homology search with a minimum of 70% identity to E. faecalis antigens listed in table 1a. The computer program TBLASTN was used to determine identity between   E.   faecalis   and   E.   faecium   sequences (http://www.hgsc.bcm.tmc.edu/microbial/microbialblast.cgi?organism=Efaecium).   The   amino   acid sequence identity was calculated based on the complete ORF sequence of the corresponding E. faecalis antigen. This result demonstrated that E. faecium contains certain structurally related proteins to the clinically closely related species E. faecalis. Since the homologous proteins in E. faecalis have been provided as hyperimmune serum reactive antigens in the present invention, the E. faecium proteins specified in Table 1b provide the same uses, especially the uses as antigens comprised in pharmaceutical composition e.g. vaccines against Enterococci infections.

Example 4: Assessment of the reactivity of highly immunogenic peptide sequences with individual human sera.

Experimental procedures

Peptide synthesis
Peptides were synthesized in small scale (4 mg resin; up to 288 in parallel) using standard F-moc chemistry on a Rink amide resin (PepChem, Tübingen, Germany) using a SyroII synthesizer (Multisyntech, Witten, Germany). After the sequence was assembled, peptides were elongated with Fmoc-epsilon-aminohexanoic acid (as a linker) and biotin (Sigma, St. Louis, MO; activated like a normal amino acid). Peptides were cleaved off the resin with 93%TFA, 5% triethylsilane, and 2% water for one hour. Peptides were dried under vacuum and freeze dried three times from acetonitrile/water (1:1). The presence of the correct mass was verified by mass spectrometry on a Reflex III MALDI-TOF (Bruker, Bremen Germany). The peptides were used without further purification.

Enzyme linked immuno assay (ELISA).
Biotin-labeled peptides (at the N-terminus) were coated on Streptavidin ELISA plates (EXICON) at 10 µg/ml concentration according to the manufacturer's instructions. Highly specific Horse Radish Peroxidase (HRP)-conjugated anti-human IgG secondary antibodies (Southern Biotech) were used according to the manufacturers' recommendations (dilution: 1,000x). Sera were tested at two serum dilutions, 200X and 1,000X. Following manual coating, peptide plates were processed and analyzed by the Gemini 160 ELISA robot (TECAN) with a built-in ELISA reader (GENIOS, TECAN).

Results

Following the bioinformatic analysis of selected clones, corresponding peptides were designed and synthesized. In case of epitopes with more than 26 amino acid residues, overlapping peptides were made. All peptides were synthesized with a N-terminal biotin-tag and used as coating reagents on Streptavidin-

coated ELISA plates.

In order to confirm the immunogenicity of selected epitopes, specific antibody levels were measured by peptide ELISA using individual human sera, which were included in the IC9- and IC10 IgG pools used for their identification in bacterial surface display screens. A summary of serum reactivity of 104 peptides present in 61 *E. faecalis* antigens with ten healthy adult screening sera (N1-N10) and eleven sera from patients with different *E. faecalis* infections are shown (P1-P11). The peptides were compared by the score calculated for each peptide based on the number of positive screening sera and the extent of reactivity. Peptides range from highly and widely reactive to weakly positive ones. Among the highest scoring peptide epitopes which are derived from surface proteins, such as LPXTG cell wall proteins (EF0490.3 and EFA0042.1-2), secreted proteins (EF0360.1 and EF0792.1) and also hypothetical proteins (EF0428.1-2, EF3207.2 and EFC0034.1-2) were found. Interestingly, two prominent epitopes belong to alternative reading frames ARF0679 (40 aa long) and ARF2052 (54 aa long) suggesting that these proteins really exist and are expressed *in vivo* during colonization and/or infection. As it was suggested by lower total *E. faecalis*-antibody levels of the patients' sera relative to the high titer healthy sera (Fig. 1C), peptide reactivities were also lower with the patients' sera.

- 55 -

Table 1a: Immunogenic proteins identified by bacterial surface display.

E. faecalis antigenic protein	Putative function (by homology)	predicted immunogenic aa*	No. of selected clones per ORF and screen	Location of identified immunogenic region (aa)	Seq. ID (DNA, Prot.)
EF0020	PTS system component	4-10,14-21,30-36,59-68,77-82,87-93,96-105,112-121,125-133,135-141,150-162,164-183,192-203,207-213,215-226,228-234,241-247,250-285,302-308	B:1, C:10, D:1	135-148	1, 171
EF0032	membrane protein, putative	15-57,60-73,77-101,108-134,136-177,185-201,203-217,226-240,244-254,272-277,283-288,292-343,354-370,380-398,406-437,439-453,473-490,532-538,584-590,595-601,606-612,664-677,679-704,715-724,731-753,759-772,786-794,814-862	A:1, C:2	657-684	2, 172
EF0062	2,3-cyclic-nucleotide 2-phosphodiesterase, putative	4-9,15-36,41-47,54-60,75-81,114-120,131-146,152-158,174-182,194-202,208-215,218-226,255-271,276-285,290-295,302-311,318-328,330-344,352-359,365-377,388-395,398-405,426-432,439-449,455-500,505-513,531-537,542-552,554-561,587-595,606-612,718-734,763-771,775-782,792-801,805-812,822-828,830-843,849-863,876-894,905-911,919-926,935-947,949-958,968-979,1009-1016,1029-1045,1047-1056,1076-1081,1092-1106,1123-1133,1179-1200,1202-1211,1215-1223,1287-1299,1301-1306	A:3, B:1, C:16, D:1	398-431, 1224-1237	3, 173
EF0149	aggregation substance ASP1	17-47,74-80,90-97,126-133,137-148,167-173,179-185,214-223,250-255,270-283,329-338,342-350,352-358,360-367,372-383,398-404,411-421,426-432,435-446,452-462,472-479,515-521,582-592,611-618,623-629,642-659,666-673,678-689,704-725,732-737,744-757,768-789,824-834,842-849,862-868,877-887,904-916,923-928,941-947,962-974,982-992,1019-1030,1032-1044,1046-1052,1065-1075,1077-1087,1108-1121,1124-1132,1137-1151,1170-1182,1190-1206,1208-1214,1227-1233,1242-1251,1254-1273,1282-1298	A:5, C:2	792-825	4, 174
EF0196	site-specific recombinase family protein	19-31,39-67,82-91,104-110,113-128,149-155,161-181	C:3	137-155	5, 175
EF0253	aldehyde dehydrogenase	6-18,54-63,69-85,110-127,142-156,158-167,169-211,238-246,248-257,276-311,339-349,371-380,385-391,394-403,421-438,451-456,483-489	C:2	449-468	6, 176
EF0270	PTS system component	5-15,24-34,50-56,61-83,98-121,123-136,149-162,166-194,202-215,221-227,229-332,337-360,367-402,404-415,427-433,444-462,471-478,487-498,511-518,521-544,550-563,568-574,580-587,597-607,610-616,624-629	A:1, B:2, C:2, D:2	468-498	7, 177
EF0298	cation-transporting ATPase, E1-E2 family	11-19,32-49,57-63,65-71,80-89,91-133,166-181,183-191,201-230,234-257,264-291,297-303,305-314,316-335,337-354,359-366,368-374,383-388,394-405,408-442,446-470,483-490,499-505,513-538,544-555,557-563,568-590,598-608,617-623,627-636,641-647,667-685,687-693,710-723,733-739,742-754,769-815	D:24	366-388	8, 178
EF0355	endolysin	4-16,30-35,42-53,67-76,82-87,101-108,112-130,132-138,147-152,161-183,187-208,218-225,265-281,295-303,305-317,322-334,338-357,360-368,370-383,387-394,400-419,421-430	A:1, B:3	255-336	9, 179
EF0428	conserved hypothetical protein	19-27,36-47,59-66,76-83,101-112,118-125,142-147,162-180,185-196,225-240,246-263,286-304,314-319,327-333,353-367	A:1, C:23	194-214	10, 180
EF0485 (EF0005)	aggregation substance	14-43,70-76,83-89,111-117,122-128,136-145,163-170,175-182,210-219,246-251,266-279,325-331,338-346,348-354,356-363,368-379,422-428,431-441,450-456,466-473,509-515,532-542,549-556,576-586,605-612,617-623,636-653,660-667,674-686,698-719,726-731,738-745,762-783,818-828,836-843,856-862,871-881,903-910,917-922,935-941,956-968,976-986,1013-1024,1026-1038,1059-1069,1071-1081,1102-1115,1118-1126,1131-1145,1164-1176,1187-1200,1202-1208,1221-1227,1236-1245,1248-1267,1273-1292	C:3, D:4	252-287, 805-844	11, 181

- 56 -

E. faecalis antigenic protein	Putative function (by homology)	predicted immunogenic aa*	No. of selected clones per ORF and screen	Location of identified immunogenic region (aa)	Seq. ID (DNA, Prot.)
EF0490	LPXTG-motif cell wall anchor domain protein	4-18,21-28,37-43,56-70,101-113,131-140,142-150,162-170,172-184,193-204,209-227,233-238,246-264	A:2, C:10	93-168	12, 182
EF0517	2-dehydropantoate 2-reductase, putative	14-20,44-50,61-70,77-96,99-106,129-142,168-181,187-196,205-221,225-241,277-296	C:4	257-281	13, 183
EF0570	Osmosensitive K+ channel sensor histidine kinase,	18-29,43-54,64-76,78-84,88-103,125-149,159-176,198-218,230-242,256-271,279-285,287-293,300-306,325-331,344-351,357-364,371-397,400-414,419-464,485-515,517-526,529-537,548-553,573-580,584-590,603-620,639-661,676-681,687-700,716-761,772-780,785-790,795-803,823-836,848-853	A:2	106-134	14, 184
EF0584	ABC transporter, ATP-binding protein	7-13,19-42,44-51,55-75,87-97,99-110,112-118,129-135,141-156,158-178,213-220,230-286,294-308,323-338,345-352,355-365,370-392,394-419,437-446,454-460,474-497,515-526,528-546,569-575	C:3	128-141	15, 185
EF0668	UDP-N-acetylmuramoylalanyl-D-glutamate-2,6-diaminopimelate ligase	12-20,24-33,45-70,73-84,86-94,103-116,118-124,135-142,163-170,176-200,202-224,226-234,237-248,250-262,265-287,296-307,334-341,347-356,361-369,382-396,405-415,418-427,431-439,443-449,452-461,467-474	D:2	113-146	16, 186
EF0792	permease, putative	13-38,44-50,52-59,66-72,83-94,103-110,116-124,131-137,158-180,199-204,218-233,241-264,269-317,326-342,350-356	C:2	70-86	17, 187
EF0795	conserved hypothetical protein	29-35,49-59,63-84,86-97,103-111,113-126,130-144,150-158,174-198,221-231,250-264,266-273,291-298,310-318	A:1, C:3	70-90	18, 188
EF0799	autolysin	19-25,28-52,60-66,71-76,131-142,149-155,157-178,181-213,218-223,237-242,250-257,260-266,272-279,282-290,321-330,373-385,393-407,441-453,461-475,509-521,529-542,577-589,597-610,643-655,663-677,703-718,729-734	A:8, B:5	358-464, 495-570, 604-685	19, 189
EF0851	hypothetical protein	4-29,51-76,116-136,158-173,179-193,207-215	C:2	86-111	20, 190
EF0861	acetyltransferase, GNAT family	5-23,45-70,79-90,93-107,114-122,142-151	C:3	18-36	21, 191
EF0922	conserved hypothetical protein	9-51,68-120,133-149,158-180,186-206,211-220,222-237,248-293,296-310,317-339	C:2	248-260	22, 192
EF0996	cell division protein FtsA	14-24,44-63,69-98,108-119,123-136,155-161,164-176,180-193,203-208,215-223,239-247,274-281,283-289,296-304,306-313,315-327,331-341,343-353,357-386,392-405	C:1, D:3	205-246	23, 193
EF1012	PTS system component	5-13,16-23,36-42,53-63,70-83,96-102	C:2	14-34	24, 194
EF1026	conserved hypothetical protein	4-13,19-35,49-56,59-76,83-107,121-134,144-153,157-164,166-186,194-202,209-216,231-253,257-264	C:21	98-134	25, 195
EF1032	daunorubicin resistance protein	16-32,38-47,58-68,78-89,98-114,117-123,132-141,146-156,164-170,179-188,196-212,219-230,232-237,244-263,265-274,278-293,297-303,306-326,339-349,352-359,362-367,373-379,384-394,396-406,423-443,451-461,465-484,490-497,504-511,523-533,537-547,550-556,558-566,573-579,586-593,598-609,615-642,647-665,671-686,693-713,723-728	A:2	332-378	26, 196
EF1060	peptide ABC transporter, peptide-binding protein, putative	6-21,34-44,58-64,66-74,79-87,114-127,129-143,154-162,174-189,205-214,241-262,266-273,278-297,319-324,328-338,342-351,390-398,409-415,422-435,458-464,471-477,481-486,506-531,534-540,542-550	A:4	315-389	27, 197
EF1093	LPXTG-motif cell wall anchor domain protein	4-28,39-45,52-58,69-82,93-115,122-128,135-140,146-163,177-192,209-215,221-232,271-284,331-337,341-352,360-378,383-390,392-401,409-422,428-435,462-470,474-480,482-496,531-539,541-549,551-560,562-569,576-582,598-618	C:17	98-127	28, 198
EF1141	MutT/nudix family protein	14-27,33-47,61-79,94-104,119-133	C:2	36-60	29, 199
EF1182	autoinducer-2	11-22,29-40,48-62,68-73,96-106,108-118,125-149	C:3	102-126	30, 200

E. faecalis antigenic protein	Putative function (by homology)	predicted immunogenic aa*	No. of selected clones per ORF and screen	Location of identified immunogenic region (aa)	Seq. ID (DNA, Prot.)
	production protein LuxS				
EF1277	transcriptional regulator	4-11,45-55,76-83,86-102,105-112,138-144,147-153	D:9	20-48	31, 201
EF1289	conserved domain protein	12-20,28-56,62-68,72-82,93-99,101-107,120-133,135-145,178-186,208-232,279-292	C:2	36-64	32, 202
EF1386	formate/nitrite transporter family protein	6-14,23-48,65-82,92-134,140-181,188-219,228-238,244-253,255-261	C:6	124-145	33, 203
EF1404	MutS2 family protein	11-25,31-38,53-59,62-71,89-99,125-133,151-157,182-190,195-203,208-215,219-229,249-262,267-275,287-295,298-316,318-325,328-334,344-353,357-363,371-377,385-391,396-415,425-436,438-457,471-485,538-552,554-561,606-625,630-636,646-653,669-679,695-704,706-715,722-747,763-773	C:6	714-738	34, 204
EF1561	shikimate 5-dehydrogenase	10-29,33-45,50-60,70-79,83-95,118-124,136-157,176-184,192-205,207-216,223-234,240-246,258-268,275-283	C:1, D:29	37-56	35, 205
EF1584	cysteine synthase A	4-24,27-38,46-54,66-72,81-97,112-119,128-137,152-157,173-179,185-214,219-225,227-248,262-284,286-295,301-307	C:6	117-134	36, 206
EF1597	catalase	26-43,49-56,60-71,74-82,87-98,110-116,131-146,154-164,169-178,183-189,205-214,241-246,255-268,275-292,305-314,316-323,326-340,346-363,397-402,419-429,440-446,452-461,467-475	C:3	29-66	37, 207
EF1601	PTS system component	7-16,21-39,48-58,61-78,82-89,109-136,138-150,152-176,182-247,255-261,267-332,336-345,347-358,362-368,371-392,394-404,407-472,490-498,505-513,527-544,554-582,603-611,614-620,632-638	A:1, B:2, C:9	500-523	38, 208
EF1624	aldehyde dehydrogenase, putative	24-46,77-83,90-97,99-118,123-166,168-177,204-212,229-239,248-262,273-282,287-293,300-319,321-337,340-352,357-366,391-402,411-428,442-450,464-471,479-489	D:2	19-40	39, 209
EF1646	heat shock protein HslVU, ATPase subunit HslU	9-23,25-34,53-58,70-86,90-97,99-116,118-128,131-141,185-191,228-233,237-253,255-261,264-271,273-280,302-312,319-349,351-359,362-369,376-383,387-394,398-406,419-434	C:3	20-31	40, 210
EF1692	hypothetical protein	15-22,37-43,71-87,105-115,121-127,135-142,152-158	D:4	32-52	41, 211
EF1741	catabolite regulator protein	6-12,18-29,37-47,50-58,65-83,85-91,94-99,108-123,142-150,156-163,183-193,215-222,242-249,252-258,261-270,285-308,318-326	A:4	1-95	42, 212
EF1798	hypothetical protein	9-61,65-133,144-155,166-173,175-221,233-276,278-313,329-368	C:2	210-233	43, 213
EF1817	serine proteinase homolog	11-29,33-39,46-51,65-93,107-113,134-143,147-154,166-177,181-188,214-220,233-243,263-269	A:2, B:1, D:3	112-128	44, 214
EF1823	autolysin	8-46,110-134,155-167,174-183,188-201,210-230,253-258,267-282,289-299,312-319,322-327,330-337,365-381,389-402,405-411,419-425,439-447,465-472,489-512,525-532,540-554,577-589,591-599,605-614,616-624,633-649	D:7	503-529	45, 215
EF1978	DNA-3-methyladenine glycosylase	34-49,64-70,90-118,124-131,141-152,159-165	C:6	112-128	46, 216
EF2005	conserved hypothetical protein	5-15,26-45,55-72,80-85,93-100,121-133,142-148,154-167,198-205,209-215,241-254,260-265,271-279	C:2	244-270	47, 217
EF2052	cell division protein FtsK	4-36,38-54,67-83,122-153,159-178,205-212,232-242,244-253,259-268,281-288,298-309,324-331,334-370,372-381,389-401,403-429,441-450,456-462,465-471,473-479,483-504,508-518,537-543,553-565,578-584,592-609,619-625,658-667,669-679,712-719,722-729,737-744,746-752,758-765	C: 6	180-226	48, 218
EF2074	ABC transporter, ATP-binding protein	6-17,23-32,49-56,61-67,76-83,85-103,105-111,120-132,145-171,175-185,191-225,231-246	C:23	99-128	49, 219
EF2266	2-deydro-3-deoxyphosphogluconate aldolase	4-24,28-48,52-58,64-79,87-100,104-120,136-152,159-166	C:4	150-163	50, 220
EF2305	toprim domain	15-27,65-71,77-99,104-121,128-154,183-216,223-229,234-255,277-	C:10	77-97	51, 221

- 58 -

E. faecalis antigenic protein	Putative function (by homology)	predicted immunogenic aa*	No. of selected clones per ORF and screen	Location of identified immunogenic region (aa)	Seq. ID (DNA, Prot.)
	protein	287,296-308			
EF2306	conserved hypothetical protein	8-18,44-76,102-109	C:2	49-57	52, 222
EF2307	conserved hypothetical protein	5-14,28-40,42-51,54-60,77-83,89-100,117-124,146-172,176-204,216-231,237-244,267-278,324-334,342-348,396-401,427-433,438-450,452-457,465-471,473-481,491-500,509-515,523-544,550-556,558-569,589-595,606-618,625-632,640-649,665-671,678-688,691-698,717-723,728-734,781-789,800-805,812-821,833-868,873-879,889-905,929-939,988-998,1046-1061,1073-1079,1089-1096,1115-1124,1132-1140,1172-1196,1220-1226,1231-1249,1269-1277,1287-1301,1307-1330,1350-1361,1369-1378,1387-1412,1414-1420,1422-1439,1484-1491,1513-1529,1552-1561,1576-1583,1606-1613,1617-1640,1647-1654,1665-1679,1686-1698,1709-1727,1736-1743,1750-1757,1771-1790,1801-1807,1817-1823,1831-1842,1859-1868,1870-1882,1884-1891,1900-1906,1909-1914,1929-1935,1952-1960,1974-1988,2002-2011,2032-2063,2071-2081,2116-2124,2139-2147,2149-2159,2163-2190,2209-2215,2245-2253,2282-2287,2331-2342,2360-2370,2379-2393,2402-2408,2414-2421,2423-2430,2433-2439,2442-2450,2472-2478,2485-2493,2495-2503,2506-2512,2547-2554,2558-2564,2615-2625,2637-2652,2692-2698,2700-2706,2711-2723,2731-2740,2748-2753,2756-2762,2765-2772,2781-2798,2810-2824,2844-2852,2885-2899,2912-2922,2937-2944,2947-2970,2988-2998,3016-3025,3032-3037,3062-3071,3129-3148,3156-3161	A: 8, B:1, C: 2,	530-607	53, 223
EF2326	reverse transcriptase	31-36,57-62,79-85,90-96,99-112,120-146,162-185,193-203,208-217,219-226,239-253,283-290,298-304,306-321,340-349,351-361,365-372,386-395,407-438,473-486,537-551,558-568,576-594,598-604	D :40	75-95	54,224
EF2378	DNA polymerase III, alpha subunit, Gram-positive type	14-19,24-30,34-42,45-52,54-64,66-82,95-105,107-118,126-163,171-177,184-201,210-215,260-269,273-279,288-304,321-327,358-364,370-375,380-387,394-404,407-413,421-431,436-451,465-474,504-511,531-552,578-587,614-626,629-636,638-671,691-715,719-729,733-745,752-759,768-777,785-792,794-802,805-824,844-854,867-880,885-891,893-902,907-924,939-948,955-964,966-975,987-1000,1012-1017,1023-1028,1050-1071,1083-1098,1102-1115,1133-1146,1170-1183,1204-1211,1213-1223,1262-1311,1313-1319,1346-1355,1366-1371,1383-1405,1409-1414	D:38	776-819	55, 225
EF2476	penicillin-binding protein 2, putative	12-27,30-38,54-61,64-74,82-96,103-110,117-125,134-140,147-158,185-201,218-225,232-253,265-280,319-325,350-362,366-372,376-386,464-483,485-490,511-521,531-537,542-559,564-574,593-609,613-619,637-642,668-677	A:3, C:8	195-214	56, 226
EF2556	succinate dehydrogenase/fumarate reductase, flavoprotein subunit	4-21,59-67,73-79,84-91,141-151,186-197,203-214,222-227,237-245,255-260,281-292,294-311,336-344,346-355,422-437,459-466,484-491	A:1, C:3	77-109	57, 227
EF2563	conserved hypothetical protein	10-45,52-61,63-70,74-102,112-122,124-132,164-178,181-205,212-240,246-256	C:2	226-247	58, 228
EF2570	xanthine dehydrogenase	38-50,53-63,78-87,89-111,126-152,169-176,179-186,193-228,254-267,275-282,288-304,309-318,325-341,346-353,358-367,384-395,404-427,429-435,456-465,467-501,510-521,523-536,541-548,552-560,563-584,589-595,597-620,625-637,639-645,661-666,712-729,734-741,743-750,775-806,809-816,818-840,842-850	C:3	693-714	59, 229
EF2581	oxidoreductase, pyridine nucleotide-disulfide	5-17,30-37,52-75,77-86,88-107,112-135,151-160,178-222,226-246,263-270,279-294,306-314,327-342,345-352,374-381,389-416,422-429,435-449,453-467,473-500,512-522,524-531,542-549,552-560,565-571,575-586,594-600,613-619,625-633,635-641,647-653,667-674,680-699,711-729,735-741,764-775,781-786,792-798,805-813,817-825,833-842,850-855,860-866,869-	A:1, C:7	533-562	60, 230

E. faecalis antigenic protein	Putative function (by homology)	predicted immunogenic aa*	No. of selected clones per ORF and screen	Location of identified immunogenic region (aa)	Seq. ID (DNA, Prot.)
		910,917-930,949-990			
EF2617	ribonuclease R	7-14,39-46,61-74,83-89,93-99,110-121,136-150,172-180,182-200,207-216,223-236,238-251,265-271,280-288,294-309,320-336,339-354,362-377,383-389,401-407,435-441,446-453,460-465,472-487,499-511,518-528,533-540,557-570,572-587,631-637,643-658,663-669,672-678,681-687,695-706,714-728	C:2	118-139	61, 231
EF2682	conserved hypothetical protein	5-19,24-30,56-64,69-79,93-100,102-111,117-123,125-133,174-182,185-199,205-224,268-275,311-336	D:20	102-125	62, 232
EF2703	transcriptional regulator	6-35,39-45,57-62,80-85,92-106,117-122,126-171,214-223,253-260,268-273,285-291,295-306,315-320,325-336,361-366	A:1, B:2, D:12	172-202	63, 233
EF2724	endoglucanase, putative	4-13,24-37,45-51,58-66,84-92,112-121,132-141,151-171,175-195,204-212,222-240,262-268,276-295,305-336,338-348,354-362	C:2	160-183	64, 234
EF2782	galactose-1-phosphate uridylyltransferase	10-16,24-35,41-73,78-104,111-121,124-139,141-148,150-164,196-215,224-241,249-282,299-307,315-357,368-378,393-401	D:2	345-367	65, 235
EF2787	rhodanese family protein	4-32,48-53,61-67,84-104,112-118	D :21	106-130	66, 236
EF2812	hypothetical protein	21-28,31-36,65-81,98-105,115-121,123-131,136-142,155-161,177-190	D :15	201-232	67, 237
EF2858	threonyl-tRNA synthetase	4-15,21-27,33-39,42-56,58-64,68-82,84-90,92-98,113-122,146-162,168-175,177-189,191-203,249-268,279-285,287-304,328-342,349-358,371-378,387-393,404-413,419-425,467-479,487-498,513-524,528-539,541-565,572-579,595-606,626-635,637-642	C:8	612-626	68, 238
EF2893	hypothetical protein	7-13,52-70,76-82,97-106,110-117	D:6	13-45	69, 239
EF2927	hydrolase, haloacid dehalogenase-like family	5-10,12-48,59-64,87-102,107-128,131-140,154-161,165-171,173-215	C:2	54-74	70, 240
EF2951	hypothetical protein	4-11,19-28,34-40,74-81,87-98,126-147,163-171,184-193,205-213	B:1, C:2	49-124	71, 241
EF2961	ribokinase	7-14,23-29,35-40,61-67,99-106,111-122,124-133,135-161,187-206,216-229,236-245,262-268,271-280	C:2	256-273	72, 242
EF2986	ABC transporter, ATP-binding protein	4-13,17-37,47-54,85-99,105-113,120-132,147-166,180-186,192-199,204-216	C:2	127-144	73, 243
EF2987	conserved hypothetical protein	14-27,29-37,52-62,68-76,89-96,117-123,125-131,137-145,166-195,205-212,214-222,228-235,258-264,271-281,288-296,308-324,332-339,355-361,365-371	A:1, C:2	268-293	74, 244
EF3023	polysaccharide lyase, family 8	4-21,30-42,54-60,78-85,90-110,141-147,160-168,176-185,194-206,218-225,230-245,251-261,287-293,295-304,320-326,334-347,351-362,386-402,413-423,427-433,439-453,456-477,480-493,507-513,526-539,574-581,591-598,600-609,614-632,655-665,685-691,703-712,742-747,757-775,797-803,813-819,823-829,880-887,901-906,930-944,948-958,962-968,971-995,1002-1009,1017-1023,1036-1053,1069-1081,1107-1124,1129-1152,1178-1195,1211-1223,1249-1266,1271-1288,1334-1340,1346-1367	A:9, D:6	1-63, 171-189	75, 245
EF3041	peptide ABC transporter, peptide-binding protein	4-22,52-63,70-75,94-104,112-125,133-141,176-199,209-216,244-259,287-299,336-352,366-372,386-399,421-436,444-449,457-466,481-487,506-529,531-540	A:2, B:1	295-378	76, 246
EF3051	hypothetical protein	9-30,43-49,58-75,86-96,119-131,138-147,162-167,181-201,208-214	A:4, C:3	16-121	77, 247
EF3060	secreted antigen SagA, putative	4-27,52-58,80-90,92-100,108-114,118-143,169-176,189-198,247-261,281-287,307-317,323-329,352-363,372-381,396-411,413-426,429-440,442-450,456-461,468-479	A:14, B:6, C:1	1-73	78, 248
EF3073	Signal peptidase I	4-32,47-52,57-63,71-78,92-104,126-142,153-175	B:3, C:3	145-163	79, 249
EF3086	conserved hypothetical protein	17-23,35-41,51-70,73-86,104-125	C:2	105-129	80, 250
EF3096	chromosome segregation SMC protein, putative	25-32,41-50,75-85,87-103,115-122,138-149,164-171,188-210,212-220,224-234,256-273,288-299,304-310,330-336,357-365,382-390,399-405,414-421,440-446,454-461,480-486,502-514,518-540,543-553,561-567,572-580,582-588,595-630,633-651,672-	C:14	909-936, 1001-1031	81, 251

E. faecalis antigenic protein	Putative function (by homology)	predicted immunogenic aa*	No. of selected clones per ORF and screen	Location of identified immunogenic region (aa)	Seq. ID (DNA, Prot.)
		681,691-709,760-767,813-832,841-848,852-866,873-893,919-925,927-933,940-955,957-978,984-997,1000-1010,1035-1040,1044-1051,1058-1064,1081-1091,1097-1124,1129-1138,1144-1150,1158-1165,1170-1180			
EF3125	primosomal protein N	4-12,19-26,31-41,49-64,66-86,101-117,119-127,134-142,152-161,163-172,179-188,209-218,234-241,276-291,294-300,307-320,324-341,346-356,373-387,389-397,410-416,418-436,444-454,460-472,481-486,500-507,511-535,541-549,553-559,579-586,602-607,613-620,628-640,654-663,671-678,681-691,709-722,741-754,766-774,778-786,797-803	C:4	212-226	82, 252
EF3177	hypothetical protein	4-10,15-27,34-54,60-73,79-88,101-115,120-136,154-162,167-172,222-240	A:3	126-195	83, 253
EF3183	LPXTG-motif cell wall anchor domain protein	5-16,18-25,29-35,57-63,86-91,107-121,123-131,170-179,185-199,204-226,250-255,262-274,291-296,325-347	A:2, B:1	1-38	84, 254
EF3207	conserved hypothetical protein	7-19,22-34,36-42,48-54,60-66,71-76,104-110,118-133,135-145,158-164,167-174,182-193,196-204,217-229,251-290,293-299,309-315	C:1, D:5	288-318	85, 255
EF3276	conserved hypothetical protein	43-51,55-61,66-73,80-90,103-127,133-142,174-180,185-196,203-210,229-235,239-251,258-266,272-278,289-314,316-326,340-346,355-361	A:1	14-27	86, 256
EF3290	sensor histidine kinase	4-25,27-33,35-41,52-74,76-89,99-124,138-144,146-159,167-182,184-191,193-206,211-223,232-240,249-257,270-279,281-287,293-310,322-341,347-356	C:14	292-322	87, 257
EF3295	hypothetical protein	5-13,28-38,43-60,67-72,98-116,122-134,137-151,167-174,177-195,197-216	B:1, C:4	99-195	88, 258
EF3319	citrate lyase, alpha subunit	15-33,35-42,48-57,62-68,73-91,107-119,121-153,173-194,205-210,223-228,234-241,243-259,275-298,308-315,327-340,342-370,376-391,398-404,410-419	A:1, C:50	71-122	89, 259
EFB0002	transposase	12-39,43-64,87-95,99-105,114-126,128-136,139-147,212-225	A:1, C:3	107-141	90, 260
EFA0041	conserved hypothetical protein	6-33,40-45,60-75,79-86,121-129,131-137,161-167,172-178,186-195,203-212,236-244,257-264,278-294,306-312,345-358,368-381,386-395,404-410,412-418	A:4, B:5, D:3	198-270	91, 261
EFA0042	LPXTG-motif cell wall anchor domain protein	18-31,34-41,50-56,69-83,99-106,129-141,147-153,159-168,170-178,190-198,200-212,221-232,237-255,261-266,274-292	A:3, B:2, C:26	118-216	92, 262
EFA0047	aggregation substance precursor	17-47,61-67,87-93,115-121,126-132,140-148,167-173,179-186,214-223,250-255,264-272,282-294,306-318,338-353,358-377,385-401,414-420,433-441,451-457,470-480,505-511,544-550,571-581,600-607,612-618,631-648,655-662,669-681,693-714,721-726,733-740,757-778,813-823,831-838,851-857,866-876,893-905,912-917,930-936,951-963,971-981,1008-1019,1021-1033,1035-1041,1054-1064,1066-1076,1097-1110,1113-1121,1126-1140,1159-1171,1182-1195,1197-1203,1216-1222,1231-1240,1243-1262,1268-1287	A:4, D:2	738-828	93, 263
EFC0015	hypothetical protein	19-28,40-46,51-57,68-74,81-87,98-108,111-121	D:3	20-36	94, 264
EFC0025	hypothetical protein	4-17,19-44,60-69,80-86,110-116	A:1, B:1, C:2	33-60	95, 265
EFC0034	hypothetical protein	8-16,18-28,42-50,53-75,79-86,94-99,122-128,136-142,149-163,166-173,198-212,254-272,288-295,304-318,324-329,343-348,351-364,367-383,389-395,411-420,427-436	C:10, D:150	11-56	96, 266
ARFC0021.1	conserved domain protein	19-25	D:5	6-24	97, 267
ARFC0021.2	conserved domain protein	6-39,59-68	D:3	44-63	98, 268
ARF0031	hypothetical protein	5-14,21-28,38-53	B:1, C:2	29-41	99, 269
ARF0066	hypothetical protein	4-13,31-41,56-65	C:14	32-56	100, 270

E. faecalis antigenic protein	Putative function (by homology)	predicted immunogenic aa*	No. of selected clones per ORF and screen	Location of identified immunogenic region (aa)	Seq. ID (DNA, Prot.)
ARF0076	hypothetical protein	5-12	A:1, C:8	4-21	101, 271
ARF0180	hypothetical protein	4-18	A :1, C :15	17-32	102, 272
ARF0275	hypothetical protein	4-10,23-33	C:2	14-30	103, 273
ARF0283	hypothetical protein	26-34,44-53	D :2	35-52	104, 274
ARF0375	hypothetical protein	none	C:3	1-19	105, 275
ARF0679	hypothetical protein	4-17,23-30,32-37	D :17	6-23	106, 276
ARF0721	hypothetical protein	5-33,40-58,61-66	C:6	45-66	107, 277
ARF1090	hypothetical protein	15-41,61-67	C:5	41-65	108, 278
ARF1583	hypothetical protein	4-12,16-23,26-37	C:4	10-29	109, 279
ARF2052	hypothetical protein	23-39	C: 37	37-55	110, 280
ARF2125	hypothetical protein	12-20	D:8	38-55	111, 281
ARF2307.1	hypothetical protein	22-37	C:4	7-22	112, 282
ARF2307.2	hypothetical protein	none	C: 2	3-14	113, 283
ARF2323	hypothetical protein	6-16,43-65,71-76	C:14	17-31	114, 284
ARF2505	hypothetical protein	4-13,27-39,42-69	B:1, C:2	17-32	115, 285
ARF2802	hypothetical protein	4-12,26-39	A:1, B:1, C:3	10-25	116, 286
ARF2902	hypothetical protein	none	C:2	2-31	117, 287
ARF3079	hypothetical protein	6-38, 49-62	D:4	39-55	118, 288
ARF3157	hypothetical protein	4-10,24-30	A:1, C:2	2-19	119, 289
ARF3182	hypothetical protein	12-17,25-46	A:2, C:2	15-30	120, 290
ARF3314	hypothetical protein	4-13	A:2, B:1, C:3, D:1	2-28	121, 291
ARFA0022	hypothetical protein	30-38	C:2	17-45	122, 292
CRF0022	hypothetical protein	24-33,55-61	C:6	31-61	123, 293
CRF0073	hypothetical protein	4-26,34-48	D:2	15-33	124, 294
CRF0096	hypothetical protein	9-15	D:3	1-22	125, 295
CRF0115	hypothetical protein	4-31	C:2	14-33	126, 296
CRF0202	hypothetical protein	5-34,49-55,64-82	C:3	69-83	127, 297
CRF0249	hypothetical protein	33-45	D:2	21-39	128, 298
CRF0258	hypothetical protein	7-14,24-32,42-65,79-86	C:3	50-64	129, 299
CRF0264	hypothetical protein	13-27,33-43,45-62	C:3	12-37	130, 300
CRF0339	hypothetical protein	4-15,17-32	C:3	10-26	131, 301
CRF0399	hypothetical protein	4-9,11-43,45-75	C:4	47-69	132, 302
CRF0682	hypothetical protein	4-18,22-37	C:12	17-34	133, 303
CRF0783	hypothetical protein	4-14	D:6	5-24	134, 304
CRF0801	hypothetical protein	7-33,35-46	C:31	1-19	135, 305
CRF0892	hypothetical protein	13-37,69-75	C:9	51-69	136, 306
CRF1041	hypothetical protein	14-24,26-34,37-49,66-78	C:5	2-25	137, 307
CRF1049	hypothetical protein	17-46,52-57,59-64	C:3	54-68	138, 308
CRF1327	hypothetical protein	4-22	C:2	13-25	139, 309
CRF1593	hypothetical protein	8-40,53-63	D:3	29-50	140, 310
CRF1610	hypothetical protein	16-28	C:2	32-40	141, 311
CRF1732	hypothetical protein	14-20,22-28,39-45	C:5, D:9	2-22	142, 312
CRF1830	hypothetical protein	4-13	C:2	12-31	143, 313
CRF1992	hypothetical protein	15-21	D:11	2-17	144, 314
CRF2074	hypothetical protein	4-17	D:2	20-36	145, 315
CRF2099	hypothetical protein	4-19	C:2	9-18	146, 316
CRF2298	hypothetical protein	4-14	D:2	3-19	147, 317
CRF2318	hypothetical protein	4-21,32-40	C:2	21-39	148, 318
CRF2568	hypothetical protein	4-13	C:2	10-27	149, 319
CRF2573	hypothetical protein	18-31,39-47,75-87,89-98	C:3	79-99	150, 320

E. faecalis antigenic protein	Putative function (by homology)	predicted immunogenic aa*	No. of selected clones per ORF and screen	Location of identified immunogenic region (aa)	Seq. ID (DNA, Prot.)
CRF2581	hypothetical protein	15-21	C:12, D:29	9-24	151, 321
CRF2647	hypothetical protein	4-14,18-27,30-53,55-64,68-74,81-98	C:7	22-40	152, 322
CRF2706	hypothetical protein	7-24,44-51	C:2	35-60	153, 323
CRF2751	hypothetical protein	10-47	C:3	23-37	154, 324
CRF2768	hypothetical protein	4-10,12-46	C:3	7-22	155, 325
CRF2778	hypothetical protein	20-27	C:3	1-13	156, 326
CRF2790	hypothetical protein	6-19,41-51	C:12	9-37	157, 327
CRF2899	hypothetical protein	4-9,11-17	C:2	9-23	158, 328
CRF2935	hypothetical protein	4-17,23-38,46-66,68-85	D:26	34-46	159, 329
CRF2966	hypothetical protein	4-18,34-59,75-81	D:31	61-84	160, 330
CRF3074	hypothetical protein	6-17	C:2	7-28	161, 331
CRF3084	hypothetical protein	4-32, 56-61	D:6	35-52	162, 332
CRF3120	hypothetical protein	4-14,27-71,74-88,93-110,115-120,124-130,139-154,161-172	C:4	146-171	163, 333
CRF3276	hypothetical protein	4-21	C:39	3-15	164, 334
CRF3277	hypothetical protein	12-17	C:11	9-26	165, 335
CRF3281	hypothetical protein	10-21,45-58	C:3	51-67	166, 336
CRF3285	hypothetical protein	59-66,68-84	D:2	13-42	167, 337
CRFC0021	hypothetical protein	11-16	C:3	1-16	168, 338
CRFC0046	hypothetical protein	4-19,23-37	C:4	10-30	169, 339
NRF0001	hypothetical protein	19-27,35-46,48-66,82-88,99-105,113-119	C:1	42-59	170, 340

Table 1b. *E.faecium* proteins identified by homology search with identified *E. faecalis* antigens

E. faecium	Putative function (by homology to E. faecalis)	E. faecalis antigenic protein	Seq.ID (DNA, Prot.)	Identity to E. faecalis (%)	Seq.ID (DNA, Prot.)
EFN0001	PTS system component	EF0020	1, 171	78	341, 357
EFN0002	conserved hypothetical protein	EF0428	10, 180	80	342, 358
EFN0003	2-dehydropantoate 2-reductase, putative	EF0517	13, 183	77	343, 359
EFN0004	conserved hypothetical protein	EF0795	18, 188	80	344, 360
EFN0005	cell division protein FtsA	EF0996	23, 193	78	345, 361
EFN0006	PTS system component	EF1012	24, 194	87	346, 362
EFN0007	LPXTG-motif cell wall anchor domain protein	EF1093	28, 198	74	347, 363
EFN0008	MutS2 family protein	EF1404	34, 204	77	348, 364
EFN0009	catabolite regulator protein	EF1741	42, 212	80	349, 365
EFN0010	2-deydro-3-deoxyphosphogluconate aldolase	EF2266	50, 220	76	350, 366
EFN0011	DNA polymerase III, alpha subunit, Gram-positive type	EF2378	55, 225	82	351, 367
EFN0012	rhodanese family protein	EF2787	66, 236	70	352, 368
EFN0013	threonyl-tRNA synthetase	EF2858	68, 238	89	353, 369
EFN0014	conserved hypothetical protein	EF3207	85, 255	85	354, 370
EFN0015	sensor histidine kinase	EF3290	87, 257	73	355, 371
EFN0016	transposase	EFB0002	90, 260	83	356, 372

Table 1c: Immunogenic proteins identified with sera from endocarditis patients by bacterial surface display.

E. faecalis antigenic protein	Putative function (by homology)	predicted immunogenic aa*	No. of selected clones per ORF and screen	Location of identified immunogenic region (aa)	Seq. ID (DNA, Prot.)
EF0008	single-strand binding protein	4-11, 16-34, 48-55, 67-77, 87-106	E:8	153-183	373, 425
EF0028	PTS system component	22-40, 49-65, 70-91, 95-109, 111-125, 146-207, 209-216, 219-225, 229-244, 251-270, 274-286, 292-309, 316-329, 335-355, 358-370, 376-388, 392-419, 425-430, 435-441, 448-455, 464-478, 486-515	F:1	437-465	374, 426
EF0146	surface exclusion protein Sea1 putative	5-19, 25-31, 43-48, 60-79, 88-100, 105-129, 148-171, 187-193, 243-263, 316-322, 334-340, 345-351, 369-378, 381-391, 399-404, 474-483, 502-517, 525-530, 558-568, 579-596, 622-627, 631-638, 644-651, 653-660, 674-680, 687-693, 721-728, 743-753, 760-775, 788-795, 806-813, 821-828, 835-842, 847-859, 868-887	F:13, H:3	300-347	375, 427
EF0153	LPXTG-motif cell wall anchor domain protein	5-26, 37-44, 89-97, 112-118, 121-128, 138-154, 157-165, 176-181, 188-198, 205-218, 223-243, 247-253, 260-279	H:1	76-155	376, 428
EF0394	44 % homology to secreted antigen SagA e.feacium	4-29, 41-46, 49-68, 82-88, 121-147, 158-164, 187-193, 195-208, 229-236, 244-249, 251-263, 269-275, 307-313, 337-343, 348-381, 392-398, 402-408, 432-438	E:96, G:2	85-117, 194-239	377, 429
EF0443	homology to LysM domain protein	5-12, 14-22, 28-34, 40-46, 70-79, 84-129, 152-165, 174-182	G:2	37-109	378, 430
EF0568	potassium-transporting ATPase subunit B	5-16, 18-52, 54-72, 81-86, 118-126, 136-145, 151-157, 168-180, 209-233, 244-270, 295-302, 315-326, 329-337, 345-352, 364-373, 397-402, 408-418, 424-431, 435-443, 472-480, 483-489, 504-510, 519-527, 549-564, 576-599, 605-637, 641-673	E:2	91-98	379, 431
EF0591	lipoprotein putative	23-36, 42-52, 133-140, 151-157, 242-247, 267-277, 295-301, 320-328, 333-339, 345-352, 365-371, 397-403, 415-428, 456-465, 481-487, 489-495, 508-516, 518-527, 585-592, 606-614, 631-637, 643-658, 665-670, 723-728, 737-744, 752-759, 787-793, 835-841, 873-885, 918-928, 938-945, 951-966, 978-988, 1015-1020, 1030-1036, 1044-1052, 1058-1069, 1071-1079, 1081-1088, 1113-1119, 1125-1138, 1141-1147, 1164-1170, 1172-1177, 1190-1200, 1214-1220, 1230-1236, 1239-1245, 1262-1268, 1270-1275, 1288-1298, 1312-1318, 1328-1334, 1337-1343, 1360-1366, 1368-1373, 1386-1396, 1410-1416, 1426-1432, 1435-1441, 1458-1464, 1484-1494, 1508-1514, 1524-1530, 1533-1539, 1556-1562	F:3	307-340	380, 432
EF0592	LPXTG-motif cell wall anchor domain protein (repeat domains)	19-25, 35-41, 44-50, 66-72, 74-79, 92-102, 116-122, 132-138, 141-147, 164-170, 172-177, 190-200, 214-220, 230-236, 239-245, 262-268, 270-275, 288-298, 312-318, 328-334, 337-343, 360-366, 368-373, 386-396, 410-416, 426-432, 435-441, 458-464, 466-478, 504-524	H:1	79-148, 177-246, 275-344, 373-442	381, 433
EF0658	hypothetical protein	7-14, 16-23, 33-39, 46-53, 72-79, 92-115, 123-130, 156-175, 179-187, 214-220, 239-246, 266-274, 302-325, 338-354, 360-370, 375-390, 392-401, 421-428, 430-463	F:2	29-58	382, 434
EF0727	conserved hypothetical protein	4-9, 22-39, 58-65, 72-82, 87-92, 99-104, 107-119, 143-166, 171-177, 194-202, 205-213, 220-228, 231-240, 247-263, 309-315, 317-323, 336-343	E:1	294-320	383, 435
EF0775	aggregation substance – chimeric	4-10, 12-18, 24-29, 34-43, 50-65, 70-76, 111-117, 129-138, 152-159, 166-171, 184-195, 200-210, 224-236, 241-251, 274-283, 285-296, 313-319, 332-341, 348-355, 378-386, 410-416, 433-445, 475-482, 523-529, 531-540, 584-596, 626-633, 674-680, 682-688, 738-750, 760-787, 828-834, 836-842, 853-862, 882-887, 893-912	H:2	604-676	384, 436
EF0779	conserved domain protein	15-38, 49-57, 60-99, 103-119, 124-194, 200-206, 215-249, 251-291, 307-313, 315-347, 369-378, 383-390, 393-400, 405-411, 423-435, 440-447, 454-460, 470-486, 490-503, 532-539, 542-	H :2	509-583	385, 437

E. faecalis antigenic protein	Putative function (by homology)	predicted immunogenic aa*	No. of selected clones per ORF and screen	Location of identified immunogenic region (aa)	Seq. ID (DNA, Prot.)
		549, 551-567, 579-592			
EF1091	hypothetical protein	38-44, 47-88, 95-103, 157-172, 235-240, 250-260, 263-276, 294-300, 312-317, 331-337, 369-391, 412-419, 442-448, 453-463, 490-529, 537-555, 571-580, 600-617, 619-627, 642-648, 682-687, 693-700, 716-722, 738-748, 756-763, 779-789, 796-802, 820-828, 833-840, 846-853, 862-872, 880-887, 894-899, 924-937, 957-963, 1006-1012, 1043-1049, 1063-1069, 1076-1097	F:4	124-147	386, 438
EF1323	conserved hypothetical protein authentic point mutation	4-28, 31-49, 60-71, 75-102, 104-114, 134-144, 160-184, 250-257, 277-285, 287-294, 330-338, 345-351, 367-374, 381-388, 393-399, 402-407, 420-426, 443-448, 458-464	F:8	411-436, 454-488	387, 439
EF1355	pyruvate dehydrogenase complex E2	20-27, 45-55, 57-64, 66-77, 98-106, 130-137, 155-165, 167-174, 176-187, 194-203, 208-223, 227-238, 245-251, 257-270, 273-278, 287-299, 330-345, 352-358, 363-385, 392-399, 410-417, 437-443, 467-484, 486-492, 495-500, 504-516, 526-536	E:2	219-270	388, 440
EF1699	transcriptional regulator MerR family	11-22, 24-31, 46-63, 65-71, 73-88, 95-109, 174-181, 183-201, 204-212, 216-222, 228-233, 241-247	H:2	142-221	389, 441
EF1744	general stress protein	8-28, 51-59, 67-84, 93-98, 140-152, 154-162, 183-188	F:3	91-125	390, 442
EF1752	conserved hypothetical protein	10-22, 27-61	F:6	69-100	391, 443
EF1753	conserved hypothetical protein	7-15, 18-26, 94-100, 126-131, 152-165, 219-228, 254-263, 274-292, 297-308, 333-340, 342-352, 354-371, 373-379, 403-410, 420-438, 450-456, 463-470, 489-495, 503-512	G:2	97-173	392, 444
EF1791	peptide ABC transporter peptide-binding protein	4-21, 37-43, 49-65, 67-74, 76-90, 113-119, 131-141, 155-173, 175-189, 192-199, 207-221, 247-254, 266-276, 317-322, 337-343, 387-393, 408-428, 439-448, 451-460, 469-479, 482-487, 493-501, 517-523, 533-542	F:1	480-503	393, 445
EF1800	conserved hypothetical protein	11-26, 40-46, 78-86, 93-103, 121-126, 132-138, 166-177, 183-196, 203-212, 214-221, 228-263, 304-311, 323-338, 345-351, 357-363, 379-393, 420-434, 442-448, 518-527, 547-553, 581-591, 602-609, 637-645, 665-674, 687-692, 701-708, 730-739, 796-802, 844-857, 882-888, 903-914, 944-950, 976-983, 1027-1033, 1049-1057, 1066-1072, 1085-1092, 1120-1127, 1137-1144, 1153-1158, 1165-1176, 1181-1187, 1221-1230, 1238-1244, 1269-1274	E:1	605-632	394, 446
EF1818	gelatinase	6-47, 57-65, 83-95, 109-121, 138-147, 154-164, 167-177, 194-200, 202-212, 227-234, 240-253, 260-267, 283-291, 320-329, 340-347, 356-364, 412-422, 430-436, 441-459, 465-475, 478-486, 498-507	F:14	59-84	395, 447
EF1850	conserved hypothetical protein	10-21, 58-83, 88-97, 120-126	F:1	21-51	396, 448
EF1877	conserved hypothetical protein	5-39, 56-62, 76-88, 90-114, 138-162, 170-195, 202-221, 228-250, 264-270, 304-355, 374-387, 391-416, 462-471, 526-546, 554-561, 574-579, 639-645, 651-660, 674-682, 689-694	E:10	666-697	397, 449
EF2174	hypothetical protein	6-30, 36-42, 143-157, 176-197, 202-209, 216-233, 241-246, 275-287, 292-299, 315-325, 343-350, 375-380, 397-403, 411-420, 422-434, 441-448, 467-474, 477-499, 555-568, 591-597, 601-609, 623-644, 667-688, 692-698, 703-718, 736-747, 757-766, 782-791, 795-801, 832-840, 859-865	F:119	226-269	398, 450
EF2224	LPXTG-motif cell wall anchor domain protein	6-23, 43-51, 61-67, 73-82, 91-97, 123-130, 149-158, 164-175, 228-234, 240-246, 248-255, 262-272, 326-332, 340-347, 365-371, 377-388, 409-419, 425-431, 438-445, 449-457, 464-470, 496-507, 559-568, 575-581, 603-608, 617-623, 633-639, 648-654, 659-670, 695-701, 734-752, 806-814, 816-829, 861-868, 891-899, 904-909, 937-945, 947-960, 978-983, 992-999, 1022-1031, 1068-1076, 1078-1091, 1109-1114, 1123-1130, 1153-1162, 1199-1207, 1209-1222, 1254-1261, 1284-1293, 1330-1338, 1340-1353, 1371-1376, 1385-1392, 1415-1421, 1433-	F:2	1422-1458	399, 451

E. faecalis antigenic protein	Putative function (by homology)	predicted immunogenic aa*	No. of selected clones per ORF and screen	Location of identified immunogenic region (aa)	Seq. ID (DNA, Prot.)
		1438, 1460-1465, 1470-1492			
EF2318	Peptidase M23/M37 family V	82-94, 111-118, 125-131, 206-212, 261-266, 310-320, 328-338, 345-351, 353-360, 414-420, 424-434, 440-447, 451-500, 506-516, 548-561, 566-572, 584-591, 601-622, 630-636, 650-659, 661-667, 674-699, 703-711, 717-729, 736-744, 752-759, 765-771, 813-822, 826-842, 852-868, 870-877, 881-895, 897-906, 913-922	H:5	602-671	400, 452
EF2704	A/G-specific adenine glycosylase	12-18, 20-25, 43-54, 56-65, 73-79, 82-88, 99-111, 136-142, 153-169, 171-183, 195-223, 229-248, 255-260, 272-277, 281-292, 298-319, 322-329, 332-351, 363-379, 381-389	F:5	275-304	401, 453
EF2713	LPXTG-motif cell wall anchor domain protein	4-9, 34-48, 65-77, 101-106, 111-131, 138-153, 186-191, 230-250	H:1	148-219	402, 454
EF2802	endolysin	4-23, 30-35, 42-53, 67-76, 82-87, 101-108, 112-130, 132-138, 147-152, 161-183, 187-208, 218-225, 265-283, 295-303, 306-317, 322-334, 338-357, 360-368, 370-383, 387-398, 400-419, 421-430	H:2	104-182, 240-304	403, 455
EF2813	conserved hypothetical protein	4-12, 63-69, 94-102, 146-164, 166-173, 175-181, 193-207, 263-281, 286-295, 301-306, 330-343, 369-378, 382-388, 414-420, 422-430, 438-454, 456-462, 472-531, 543-560, 581-591, 596-605, 614-623, 626-635, 656-662, 669-676, 683-690, 693-698, 705-711, 728-736, 752-764	E:4	69-102	404, 456
EF2820	hypothetical protein	6-12, 43-53, 141-147, 164-179, 185-195, 197-206, 227-235, 237-271, 288-305, 308-317, 335-341, 351-357, 365-376, 386-395, 397-416, 422-447	E:1	11-35	405, 457
EF3082	iron compound ABC transporter substrate-binding protein	16-24, 50-65, 73-84, 88-99, 114-124, 130-146, 181-187, 193-203, 214-220, 236-247, 250-258, 287-297	F:92	50-113	406, 458
EF3256	lipoprotein putative	4-25, 50-55, 76-82, 117-123, 131-137, 139-148, 157-166, 239-245, 253-258, 266-275, 277-292, 300-306	F:5, G:3	51-83, 93-161	407, 459
EFA0021	hypothetical protein	6-22, 34-43, 51-86, 93-100, 110-116, 150-161, 164-171, 180-187, 197-218	H:2	168-237	408, 460
EFA0044	hypothetical protein	4-27, 55-60, 74-82	E:6	10-46	409, 461
EFA0052	surface exclusion protein Sea1	6-19, 25-31, 43-49, 60-79, 88-100, 105-129, 148-161, 164-171, 187-193, 243-263, 316-322, 334-340, 369-378, 381-391, 398-404, 460-466, 474-483, 502-509, 511-517, 525-530, 558-567, 579-596, 622-627, 631-638, 641-651, 653-659, 674-680, 687-693, 710-716, 720-727, 743-753, 760-775, 788-795, 806-813, 821-828, 836-842, 847-860, 865-880	H:10	258-377	410, 462
EFC0004	TraC protein	4-11, 25-64, 71-79, 88-94, 107-120, 123-132, 167-188, 231-237, 240-246, 261-267, 306-311, 330-342, 351-358, 389-395, 406-418, 429-434, 439-448, 483-501, 511-520	G:1	71-143	411, 463
EFC0012	LPXTG-motif cell wall anchor domain protein	4-18, 22-27, 53-64, 94-100, 121-127, 133-139, 155-164, 177-182, 187-196, 206-218, 224-242, 248-253, 258-277	H:1	184-253	412, 464
EFC0021	conserved domain protein	10-17, 56-67, 72-82, 94-99, 106-113, 166-173, 229-235, 243-283, 295-301, 313-321, 326-331, 342-348, 396-414, 423-435, 446-452, 454-462, 496-502, 511-534, 543-556, 563-570, 586-593, 616-626, 638-645, 653-662, 679-696, 731-737, 766-774, 776-782, 790-796, 810-817, 825-835, 837-846	H:5	540-615	413, 465
EFC0053	Transposase Mutator family	13-24, 30-36, 73-81, 89-95, 109-115, 117-143, 161-173, 179-189, 226-244, 251-261, 275-281, 298-305, 307-315, 323-328, 364-374	G:6, H:12	69-186, 264-354	414, 466
ARFC0021.1	hypothetical protein	19-25	D:5	6-22	415, 467
ARFC0021.2	hypothetical protein	6-39, 59-68	D:3	43-62	416, 468
ARF0324	hypothetical protein	6-14, 22-32	F:2	1-27	417, 469
ARF1627	hypothetical protein	4-41	E:127	28-40	418, 470
ARF1650	hypothetical protein	8-14	F:5	4-19	419, 471
CRF0097	hypothetical protein	4-10, 12-22, 30-35	E:78	6-33	420, 472

- 67 -

E. faecalis antigenic protein	Putative function (by homology)	predicted immunogenic aa*	No. of selected clones per ORF and screen	Location of identified immunogenic region (aa)	Seq. ID (DNA, Prot.)
CRF0257	hypothetical protein	4-16, 24-33	F:21	37-54	421, 473
CRF0635	hypothetical protein	none	E:7	2-23	422, 474
CRF1152	hypothetical protein	4-21, 27-33, 36-41	E:16	14-34	423, 475
CRF1720	hypothetical protein	4-14, 24-30, 37-42, 57-78, 83-89, 94-103, 113-131	E:5	100-122	424, 476

Table 2: Epitope serology with human sera.

Peptide	Seq. ID	location in protein (aa)	S
EF0020.1	171	135-147	12
EF0032.1	172	658-682	15
EF0062.1	173	411-427	16
EF0062.2	173	1226-1246	7
EF0149.1	174	794-817	11
EF0149.2	174	801-824	10
EF0270.1	177	468-492	24
EF0270.2	177	474-495	7
EF0298.1	178	366-388	12
EF0355.1	179	266-291	8
EF0355.2	179	287-312	8
EF0355.3	179	308-333	5
EF0428.1	180	197-213	26
EF0428.2	180	195-211	31
EF0485.1	181	252-275	8
EF0485.2	181	262-285	12
EF0485.3	181	812-830	8
EF0490.1	182	94-112	9
EF0490.2	182	97-120	15
EF0490.3	182	104-128	20
EF0517.1	183	257-281	14
EF0570.1	184	106-134	14
EF0792.1	187	70-86	28
EF0799.1	189	358-383	8
EF0799.2	189	378-402	5
EF0799.3	189	397-421	5
EF0799.4	189	499-524	10
EF0799.5	189	520-545	8
EF0799.6	189	541-566	9
EF0799.7	189	622-646	4
EF0799.8	189	641-665	7
EF0799.9	189	660-684	8
EF0922.1	192	248-260	15
EF1012.1	194	15-34	4
EF1026.1	195	112-129	10
EF1032.1	196	333-358	12
EF1032.2	196	353-378	3
EF1060.1	197	316-343	7
EF1060.2	197	339-366	8
EF1060.3	197	362-389	16
EF1093.1	198	98-123	11
EF1093.2	198	104-126	7
EF1277.1	201	20-43	15
EF1277.2	201	23-48	15
EF1386.1	203	124-145	6
EF1404.1	204	717-738	4
EF1561.1	205	37-56	4
EF1584.1	206	118-134	17
EF1601.1	208	500-522	5
EF1692.1	211	32-47	16
EF1741.1	212	25-51	10
EF1741.2	212	47-73	7
EF1741.3	212	69-95	11
EF1823.1	215	503-529	11

(Column headers for the serology grid: N1, N2, N3, N4, N5, N6, N7, N8, N9, N10, P1, P2, P3, P4, P5, P6, P7, P8, P9, P10, P11)

EF1978.1	216	112-128	14
EF2052.1	218	181-199	8
EF2074.1	219	109-121	17
EF2266.1	220	150-163	12
EF2305.1	221	77-97	10
EF2307.1	223	564-586	4
EF2326.1	224	75-94	4
EF2378.1	225	776-798	4
EF2378.2	225	784-808	11
EF2378.3	225	794-815	20
EF2476.1	226	196-212	14
EF2556.1	226	78-100	3
EF2556.2	226	85-107	4
EF2581.1	230	536-553	14
EF2682.1	232	102-125	7
EF2703.1	233	178-198	24
EF2858.1	238	612-626	13
EF3023.1	245	171-187	14
EF3041.1	246	296-320	1
EF3041.2	246	315-339	4
EF3041.3	246	334-358	3
EF3041.4	246	353-377	7
EF3051.1	247	47-71	16
EF3060.1	248	1-25	28
EF3060.2	248	20-45	5
EF3060.3	248	40-64	1
EF3073.1	249	146-161	16
EF3096.1	251	910-935	15
EF3096.2	251	1007-1030	7
EF3125.1	252	212-226	14
EF3177.1	253	126-152	19
EF3177.2	253	148-173	7
EF3177.3	253	169-195	7
EF3207.1	255	288-310	2
EF3207.2	255	293-316	16
EF3290.1	257	293-312	3
EF3295.1	258	154-170	19
EF3319.1	259	72-95	7
EF3319.2	259	90-112	7
EF3319.3	259	97-121	15
EFA0042.1	262	135-150	31
EFA0042.2	262	146-163	23
EFA0047.1	263	799-827	10
EFC0034.1	266	23-43	16
EFC0034.2	266	33-53	16
ARFC0021.1	268	44-62	6
ARF0679.1	276	6-22	36
ARF2052.1	280	37-54	29
ARF2125.1	281	40-54	17
ARF2307.1	282	7-21	15
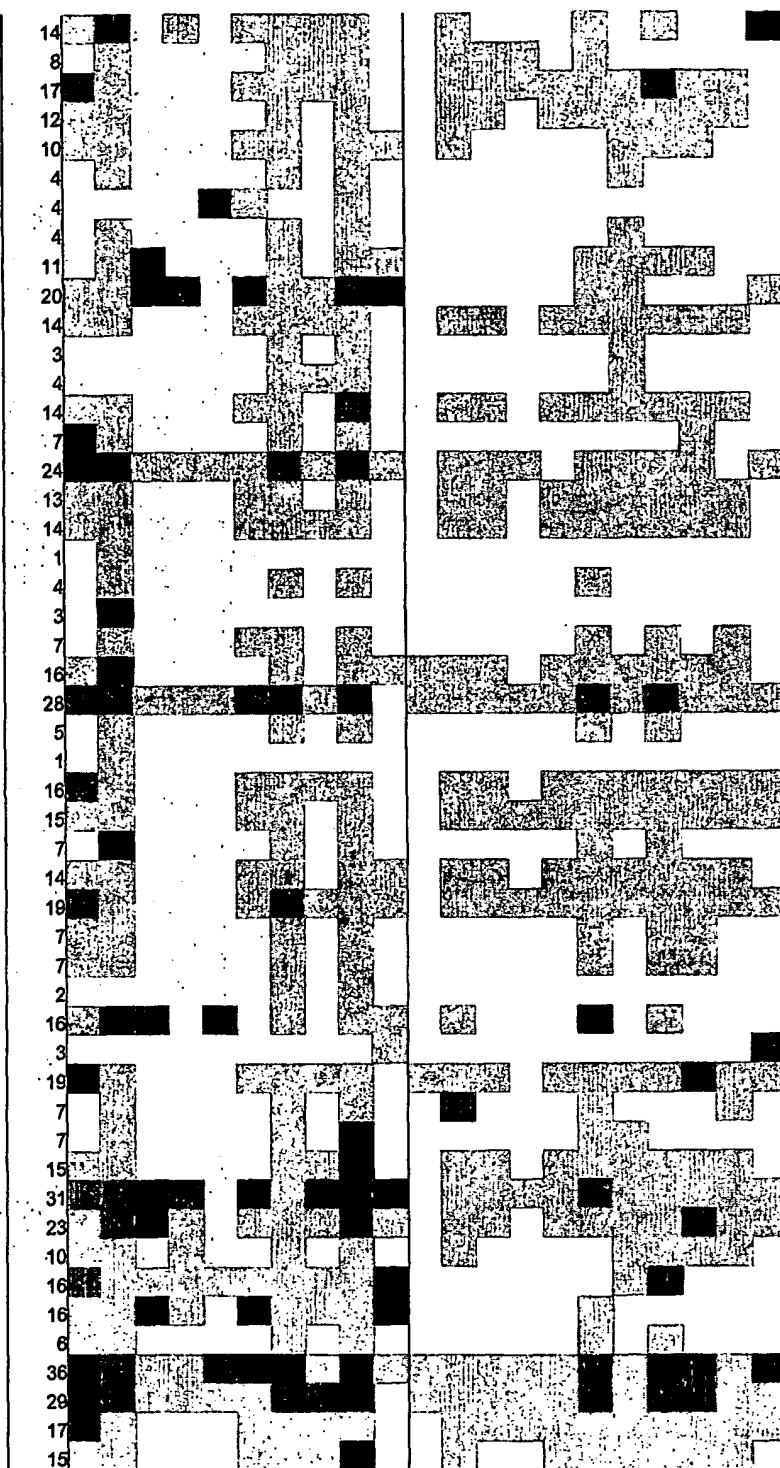
References

Altschul, S., et al. (1990). Journal of Molecular Biology 215: 403-10.
Bennett, D., et al. (1995). J Mol Recognit 8: 52-8.
Burnie, J., et al. (1998). J Antimicrob Chemother 41: 319-22.
Cetinkaya, Y., et al. (2000). Clin Microbiol Rev 13: 686-707.
Clackson, T., et al. (1991). Nature 352: 624-8.
Devereux, J., et al. (1984). Nucleic acids research 12: 387-95.
Doherty, E., et al. (2001). Annu Rev Biophys Biomol Struct 30: 457-475.
Eisenbraun, M., et al. (1993). DNA Cell Biol 12: 791-7.
Etz, H., et al. (2001). J Bacteriol 183: 6924-35.
Ferretti, J., et al. (1986). J Bacteriol 167: 631-8.
French, G. (1998). Clin Infect Dis 27: S75-83.
Gaglani, M., et al. (1997). J Clin Immunol 17: 478-84.
Ganz, T. (1999). Science 286: 420-421.
Georgiou, G. (1997). Nature Biotechnology 15: 29-34.
Gold, H. (2001). Clin Infect Dis 33: 210-9.
Haas, W., et al. (2002). Nature 415: 84-7.
Hancock, L.E., et al. (2000) pp251-258. In Fischetti, V.A., et al. (ed.), Gram-positive pathogens. AMS Press.
Hashemzadeh-Bonehi, L., et al. (1998). Mol Microbiol 30: 676-678.
Hemmer, B., et al. (1999). Nat Med 5: 1375-82.
Hoe, N., et al. (2001). J Infect Dis 183: 633-9.
Huebner, J., et al. (2000). Infect Immun 68: 4631-6.
Jett, B., et al. (1994). Clin Microbiol Rev 7: 462-78.
Johanson, K., et al. (1995). J Biol Chem 270: 9459-71.
Jones, P., et al. (1986). Nature 321: 522-5.
Kajava, A., et al. (2000). J Bacteriol 182: 2163-9.
Kohler, G., et al. (1975). Nature 256: 495-7.
Lewin, A., et al. (2001). Trends Mol Med 7: 221-8.
Lowe, A., et al. (1995). Infect Immun 63: 703-6.
Marks, J., et al. (1992). Biotechnology (N Y) 10: 779-83.
McCafferty, J., et al. (1990). Nature 348: 552-4.
McCormick, J., et al. (2001). Infect Immun 69: 3305-14.
Murray, B. (1990). Clin Microbiol Rev 3: 46-65.
Navarre, W., et al. (1999). Microbiol Mol Biol Rev 63: 174-229.
Noble, W., et al. (1992). FEMS Microbiol Lett 72: 195-8.
Okano, H., et al. (1991). J Neurochem 56: 560-7.
Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression; CRC Press, Boca Ration, FL (1988) for a description of these molecules.
Paulsen, I., et al. (2003). Science 299: 2071-4.
Poyart, C., et al. (1997). Antimicrob Agents Chemother 41: 24-9.
Rammensee, H., et al. (1999). Immunogenetics 50: 213-9.
Rice, L. (2001). Emerg Infect Dis 7: 183-7.
Richards, M., et al. (2000). Infect Control Hosp Epidemiol 21: 510-5.
Rosenshine, I., et al. (1992). Infect Immun 60: 2211-7.
Seeger, C., et al. (1984). Proc Natl Acad Sci U S A 81: 5849-52.
Shankar, V., et al. (1999). Infect Immun 67: 193-200.
Skerra, A. (1994). Gene 151: 131-5.
Sussmuth, S., et al. (2000). Infect Immun 68: 4900-6.
Tang, D., et al. (1992). Nature 356: 152-4.
Tempest, P., et al. (1991). Biotechnology (N Y) 9: 266-71.
Tourdot, S., et al. (2000). Eur J Immunol 30: 3411-21.

Whitnack, E., et al. (1985). <u>J Exp Med</u> 162: 1983-97.
Whiley, J., et al. (1987). Current Protocols in Molecular Biology.
Xu, Y., et al. (1997). <u>Infect Immun</u> 65: 4207-15.

Claims:

1.   An isolated nucleic acid molecule encoding a hyperimmune serum reactive antigen or a fragment thereof comprising a nucleic acid sequence which is selected from the group consisting of:
     a)   a nucleic acid molecule having at least 70% sequence identity to a nucleic acid molecule selected from Seq ID No 1-2, 4-8, 10, 12-18, 20-23, 25-26, 29-43, 45-62, 64-74, 76-77, 79-83, 85-89, 91-92, 94-114, 117-126, 128-146, 148-170, 373, 375, 379-381, 387, 392, 394, 397-399, 407-408, 410-411 and 415-424,
     b)   a nucleic acid molecule which is complementary to the nucleic acid molecule of a),
     c)   a nucleic acid molecule comprising at least 15 sequential bases of the nucleic acid molecule of a) or b)
     d)   a nucleic acid molecule which anneals under stringent hybridisation conditions to the nucleic acid molecule of a), b), or c)
     e)   a nucleic acid molecule which, but for the degeneracy of the genetic code, would hybridise to the nucleic acid molecule defined in a), b), c) or d).

2.   The isolated nucleic acid molecule according to claim 1, wherein the sequence identity is at least 80%, preferably at least 95%, especially 100%.

3.   An isolated nucleic acid molecule encoding a hyperimmune serum reactive antigen or a fragment thereof comprising a nucleic acid sequence selected from the group consisting of
     a)   a nucleic acid molecule having at least 96%, preferably at least 98 %, especially 100 % sequence identity to a nucleic acid molecule selected from Seq ID No 3, 9, 11, 24, 27, 44, 63, 75, 84, 115-116, 127, 374, 376-378, 382-386, 388-391, 393, 395-396, 400-406, 409 and 412-414.
     b)   a nucleic acid molecule which is complementary to the nucleic acid molecule of a),
     c)   a nucleic acid molecule comprising at least 15 sequential bases of the nucleic acid molecule of a) or b)
     d)   a nucleic acid molecule which anneals under stringent hybridisation conditions to the nucleic acid molecule of a), b) or c),
     e)   a nucleic acid molecule which, but for the degeneracy of the genetic code, would hybridise to the nucleic acid defined in a), b), c) or d).

4.   An isolated nucleic acid molecule comprising a nucleic acid sequence selected from the group consisting of
     a)   a nucleic acid molecule selected from Seq ID No 90, 147,
     b)   a nucleic acid molecule which is complementary to the nucleic acid of a),
     c)   a nucleic acid molecule which, but for the degeneracy of the genetic code, would hybridise to the nucleic acid defined in a), b), c) or d).

5.   The nucleic acid molecule according to any one of the claims 1, 2, 3 or 4, wherein the nucleic acid is DNA.

6.   The nucleic acid molecule according to any one of the claims 1,2, 3, 4, or 5 wherein the nucleic acid is RNA.

7.   An isolated nucleic acid molecule according to any one of claims 1 to 5, wherein the nucleic acid molecule is isolated from a genomic DNA, especially from a *E. faecalis* genomic DNA.

8.   A vector comprising a nucleic acid molecule according to any one of claims 1 to 7.

9.   A vector according to claim 8, wherein the vector is adapted for recombinant expression of the hyperimmune serum reactive antigens or fragment thereof encoded by the nucleic acid molecule

according to any one of claims 1 to 7.

10. A host cell comprising the vector according to claim 8 or 9.

11. A hyperimmune serum-reactive antigen comprising an amino acid sequence being encoded by a nucleic acid molecule according to any one of the claims 1, 2, 5, 6 or 7 and fragments thereof, wherein the amino acid sequence is selected from the group consisting of Seq ID No 171-172, 174-178, 180, 182-188, 190-193, 195-196, 199-213, 215-232, 234-244, 246-247, 249-253, 255-259, 261-262, 264-284, 287-296, 298-316, 318-340, 425, 427, 431-433, 439, 444, 446, 449-451, 459-460, 462-463 and 467-476.

12. A hyperimmune serum-reactive antigen comprising an amino acid sequence being encoded by a nucleic acid molecule according to any one of the claims 3, 5, 6, or 7 and fragments thereof, wherein the amino acid sequence is selected from the group consisting of Seq ID No 173, 179, 181, 194, 197, 214, 233, 245, 254, 285-286, 297, 426, 428-430, 434-438, 440-443, 445, 447-448, 452-458, 461 and 464-466.

13. A hyperimmune serum-reactive antigen comprising an amino acid sequence being encoded by a nucleic acid molecule according to any one of the claims 4, 5, 6, or 7 and fragments thereof, wherein the amino acid sequence is selected from the group consisting of Seq ID No 260, 317.

14. Fragments of hyperimmune serum-reactive antigens selected from the group consisting of peptides comprising amino acid sequences of column "predicted immunogenic aa" and "location of identified immunogenic region" of Table 1a and Table 1c; the serum reactive epitopes of Table 2, especially peptides comprising amino acid 4-10, 14-21, 30-36, 59-68, 77-82, 87-93, 96-105, 112-121, 125-133, 135-141, 150-162, 164-183, 192-203, 207-213, 215-226, 228-234, 241-247, 250-285, 302-308 and 135-148 of Seq ID No 171; 15-57, 60-73, 77-101, 108-134, 136-177, 185-201, 203-217, 226-240, 244-254, 272-277, 283-288, 292-343, 354-370, 380-398, 406-437, 439-453, 473-490, 532-538, 584-590, 595-601, 606-612, 664-677, 679-704, 715-724, 731-753, 759-772, 786-794, 814-862 and 657-684 of Seq ID No 172; 4-9, 15-36, 41-47, 54-60, 75-81, 114-120, 131-146, 152-158, 174-182, 194-202, 208-215, 218-226, 255-271, 276-285, 290-295, 302-311, 318-328, 330-344, 352-359, 365-377, 388-395, 398-405, 426-432, 439-449, 455-500, 505-513, 531-537, 542-552, 554-561, 587-595, 606-612, 718-734, 763-771, 775-782, 792-801, 805-812, 822-828, 830-843, 849-863, 876-894, 905-911, 919-926, 935-947, 949-958, 968-979, 1009-1016, 1029-1045, 1047-1056, 1076-1081, 1092-1106, 1123-1133, 1179-1200, 1202-1211, 1215-1223, 1287-1299, 1301-1306, 398-431 and 1224-1237 of Seq ID No 173; 17-47, 74-80, 90-97, 126-133, 137-148, 167-173, 179-185, 214-223, 250-255, 270-283, 329-338, 342-350, 352-358, 360-367, 372-383, 398-404, 411-421, 426-432, 435-446, 452-462, 472-479, 515-521, 582-592, 611-618, 623-629, 642-659, 666-673, 678-689, 704-725, 732-737, 744-757, 768-789, 824-834, 842-849, 862-868, 877-887, 904-916, 923-928, 941-947, 962-974, 982-992, 1019-1030, 1032-1044, 1046-1052, 1065-1075, 1077-1087, 1108-1121, 1124-1132, 1137-1151, 1170-1182, 1190-1206, 1208-1214, 1227-1233, 1242-1251, 1254-1273, 1282-1298 and 792-825 of Seq ID No 174; 19-31, 39-67, 82-91, 104-110, 113-128, 149-155, 161-181 and 137-155 of Seq ID No 175; 6-18, 54-63, 69-85, 110-127, 142-156, 158-167, 169-211, 238-246, 248-257, 276-311, 339-349, 371-380, 385-391, 394-403, 421-438, 451-456, 483-489 and 449-468 of Seq ID No 176; 5-15, 24-34, 50-56, 61-83, 98-121, 123-136, 149-162, 166-194, 202-215, 221-227, 229-332, 337-360, 367-402, 404-415, 427-433, 444-462, 471-478, 487-498, 511-518, 521-544, 550-563, 568-574, 580-587, 597-607, 610-616, 624-629 and 468-498 of Seq ID No 177; 11-19, 32-49, 57-63, 65-71, 80-89, 91-133, 166-181, 183-191, 201-230, 234-257, 264-291, 297-303, 305-314, 316-335, 337-354, 359-366, 368-374, 383-388, 394-405, 408-442, 446-470, 483-490, 499-505, 513-538, 544-555, 557-563, 568-590, 598-608, 617-623, 627-636, 641-647, 667-685, 687-693, 710-723, 733-739, 742-754, 769-815 and 366-388 of Seq ID No 178; 4-16, 30-35, 42-53, 67-76, 82-87, 101-108, 112-130, 132-138, 147-152, 161-183, 187-208, 218-225, 265-281, 295-303, 305-317, 322-334, 338-357, 360-368, 370-383, 387-394, 400-419, 421-430 and 255-336 of Seq ID No 179; 19-

27, 36-47, 59-66, 76-83, 101-112, 118-125, 142-147, 162-180, 185-196, 225-240, 246-263, 286-304, 314-319, 327-333, 353-367 and 194-214 of Seq ID No 180; 14-43, 70-76, 83-89, 111-117, 122-128, 136-145, 163-170, 175-182, 210-219, 246-251, 266-279, 325-331, 338-346, 348-354, 356-363, 368-379, 422-428, 431-441, 450-456, 466-473, 509-515, 532-542, 549-556, 576-586, 605-612, 617-623, 636-653, 660-667, 674-686, 698-719, 726-731, 738-745, 762-783, 818-828, 836-843, 856-862, 871-881, 903-910, 917-922, 935-941, 956-968, 976-986, 1013-1024, 1026-1038, 1059-1069, 1071-1081, 1102-1115, 1118-1126, 1131-1145, 1164-1176, 1187-1200, 1202-1208, 1221-1227, 1236-1245, 1248-1267, 1273-1292, 252-287 and 805-844 of Seq ID No 181; 4-18, 21-28, 37-43, 56-70, 101-113, 131-140, 142-150, 162-170, 172-184, 193-204, 209-227, 233-238, 246-264 and 93-168 of Seq ID No 182; 14-20, 44-50, 61-70, 77-96, 99-106, 129-142, 168-181, 187-196, 205-221, 225-241, 277-296 and 257-281 of Seq ID No 183; 18-29, 43-54, 64-76, 78-84, 88-103, 125-149, 159-176, 198-218, 230-242, 256-271, 279-285, 287-293, 300-306, 325-331, 344-351, 357-364, 371-397, 400-414, 419-464, 485-515, 517-526, 529-537, 548-553, 573-580, 584-590, 603-620, 639-661, 676-681, 687-700, 716-761, 772-780, 785-790, 795-803, 823-836, 848-853 and 106-134 of Seq ID No 184; 7-13, 19-42, 44-51, 55-75, 87-97, 99-110, 112-118, 129-135, 141-156, 158-178, 213-220, 230-286, 294-308, 323-338, 345-352, 355-365, 370-392, 394-419, 437-446, 454-460, 474-497, 515-526, 528-546, 569-575 and 128-141 of Seq ID No 185; 12-20, 24-33, 45-70, 73-84, 86-94, 103-116, 118-124, 135-142, 163-170, 176-200, 202-224, 226-234, 237-248, 250-262, 265-287, 296-307, 334-341, 347-356, 361-369, 382-396, 405-415, 418-427, 431-439, 443-449, 452-461, 467-474 and 113-146 of Seq ID No 186; 13-38, 44-50, 52-59, 66-72, 83-94, 103-110, 116-124, 131-137, 158-180, 199-204, 218-233, 241-264, 269-317, 326-342, 350-356 and 70-86 of Seq ID No 187; 29-35, 49-59, 63-84, 86-97, 103-111, 113-126, 130-144, 150-158, 174-198, 221-231, 250-264, 266-273, 291-298, 310-318 and 70-90 of Seq ID No 188; 19-25, 28-52, 60-66, 71-76, 131-142, 149-155, 157-178, 181-213, 218-223, 237-242, 250-257, 260-266, 272-279, 282-290, 321-330, 373-385, 393-407, 441-453, 461-475, 509-521, 529-542, 577-589, 597-610, 643-655, 663-677, 703-718, 729-734, 358-464, 495-570 and 604-685 of Seq ID No 189; 4-29, 51-76, 116-136, 158-173, 179-193, 207-215 and 86-111 of Seq ID No 190; 5-23, 45-70, 79-90, 93-107, 114-122, 142-151 and 18-36 of Seq ID No 191; 9-51, 68-120, 133-149, 158-180, 186-206, 211-220, 222-237, 248-293, 296-310, 317-339 and 248-260 of Seq ID No 192; 14-24, 44-63, 69-98, 108-119, 123-136, 155-161, 164-176, 180-193, 203-208, 215-223, 239-247, 274-281, 283-289, 296-304, 306-313, 315-327, 331-341, 343-353, 357-386, 392-405 and 205-246 of Seq ID No 193; 5-13, 16-23, 36-42, 53-63, 70-83, 96-102 and 14-34 of Seq ID No 194; 4-13, 19-35, 49-56, 59-76, 83-107, 121-134, 144-153, 157-164, 166-186, 194-202, 209-216, 231-253, 257-264 and 98-134 of Seq ID No 195; 16-32, 38-47, 58-68, 78-89, 98-114, 117-123, 132-141, 146-156, 164-170, 179-188, 196-212, 219-230, 232-237, 244-263, 265-274, 278-293, 297-303, 306-326, 339-349, 352-359, 362-367, 373-379, 384-394, 396-406, 423-443, 451-461, 465-484, 490-497, 504-511, 523-533, 537-547, 550-556, 558-566, 573-579, 586-593, 598-609, 615-642, 647-665, 671-686, 693-713, 723-728 and 332-378 of Seq ID No 196; 6-21, 34-44, 58-64, 66-74, 79-87, 114-127, 129-143, 154-162, 174-189, 205-214, 241-262, 266-273, 278-297, 319-324, 328-338, 342-351, 390-398, 409-415, 422-435, 458-464, 471-477, 481-486, 506-531, 534-540, 542-550 and 315-389 of Seq ID No 197; 4-28, 39-45, 52-58, 69-82, 93-115, 122-128, 135-140, 146-163, 177-192, 209-215, 221-232, 271-284, 331-337, 341-352, 360-378, 383-390, 392-401, 409-422, 428-435, 462-470, 474-480, 482-496, 531-539, 541-549, 551-560, 562-569, 576-582, 598-618 and 98-127 of Seq ID No 198; 14-27, 33-47, 61-79, 94-104, 119-133 and 36-60 of Seq ID No 199; 11-22, 29-40, 48-62, 68-73, 96-106, 108-118, 125-149 and 102-126 of Seq ID No 200; 4-11, 45-55, 76-83, 86-102, 105-112, 138-144, 147-153 and 20-48 of Seq ID No 201; 12-20, 28-56, 62-68, 72-82, 93-99, 101-107, 120-133, 135-145, 178-186, 208-232, 279-292 and 36-64 of Seq ID No 202; 6-14, 23-48, 65-82, 92-134, 140-181, 188-219, 228-238, 244-253, 255-261 and 124-145 of Seq ID No 203; 11-25, 31-38, 53-59, 62-71, 89-99, 125-133, 151-157, 182-190, 195-203, 208-215, 219-229, 249-262, 267-275, 287-295, 298-316, 318-325, 328-334, 344-353, 357-363, 371-377, 385-391, 396-415, 425-436, 438-457, 471-485, 538-552, 554-561, 606-625, 630-636, 646-653, 669-679, 695-704, 706-715, 722-747, 763-773 and 714-738 of Seq ID No 204; 10-29, 33-45, 50-60, 70-79, 83-95, 118-124, 136-157, 176-184, 192-205, 207-216, 223-234, 240-246, 258-268, 275-283 and 37-56 of Seq ID No 205; 4-24, 27-38, 46-54, 66-72, 81-97, 112-119, 128-137, 152-157, 173-179, 185-214, 219-225, 227-248, 262-284, 286-295, 301-307 and 117-134 of Seq ID No 206; 26-43, 49-56, 60-71, 74-82, 87-98, 110-116, 131-146, 154-164, 169-178, 183-189, 205-214, 241-246, 255-268, 275-292, 305-314, 316-323, 326-340, 346-363, 397-402, 419-429, 440-446, 452-461, 467-475 and 29-66 of Seq ID No 207; 7-16,

21-39, 48-58, 61-78, 82-89, 109-136, 138-150, 152-176, 182-247, 255-261, 267-332, 336-345, 347-358, 362-368, 371-392, 394-404, 407-472, 490-498, 505-513, 527-544, 554-582, 603-611, 614-620, 632-638 and 500-523 of Seq ID No 208; 24-46, 77-83, 90-97, 99-118, 123-166, 168-177, 204-212, 229-239, 248-262, 273-282, 287-293, 300-319, 321-337, 340-352, 357-366, 391-402, 411-428, 442-450, 464-471, 479-489 and 19-40 of Seq ID No 209; 9-23, 25-34, 53-58, 70-86, 90-97, 99-116, 118-128, 131-141, 185-191, 228-233, 237-253, 255-261, 264-271, 273-280, 302-312, 319-349, 351-359, 362-369, 376-383, 387-394, 398-406, 419-434 and 20-31 of Seq ID No 210; 15-22, 37-43, 71-87, 105-115, 121-127, 135-142, 152-158 and 32-52 of Seq ID No 211; 6-12, 18-29, 37-47, 50-58, 65-83, 85-91, 94-99, 108-123, 142-150, 156-163, 183-193, 215-222, 242-249, 252-258, 261-270, 285-308, 318-326 and 1-95 of Seq ID No 212; 9-61, 65-133, 144-155, 166-173, 175-221, 233-276, 278-313, 329-368 and 210-233 of Seq ID No 213; 11-29, 33-39, 46-51, 65-93, 107-113, 134-143, 147-154, 166-177, 181-188, 214-220, 233-243, 263-269 and 112-128 of Seq ID No 214; 8-46, 110-134, 155-167, 174-183, 188-201, 210-230, 253-258, 267-282, 289-299, 312-319, 322-327, 330-337, 365-381, 389-402, 405-411, 419-425, 439-447, 465-472, 489-512, 525-532, 540-554, 577-589, 591-599, 605-614, 616-624, 633-649 and 503-529 of Seq ID No 215; 34-49, 64-70, 90-118, 124-131, 141-152, 159-165 and 112-128 of Seq ID No 216; 5-15, 26-45, 55-72, 80-85, 93-100, 121-133, 142-148, 154-167, 198-205, 209-215, 241-254, 260-265, 271-279 and 244-270 of Seq ID No 217; 4-36, 38-54, 67-83, 122-153, 159-178, 205-212, 232-242, 244-253, 259-268, 281-288, 298-309, 324-331, 334-370, 372-381, 389-401, 403-429, 441-450, 456-462, 465-471, 473-479, 483-504, 508-518, 537-543, 553-565, 578-584, 592-609, 619-625, 658-667, 669-679, 712-719, 722-729, 737-744, 746-752, 758-765 and 180-226 of Seq ID No 218; 6-17, 23-32, 49-56, 61-67, 76-83, 85-103, 105-111, 120-132, 145-171, 175-185, 191-225, 231-246 and 99-128 of Seq ID No 219; 4-24, 28-48, 52-58, 64-79, 87-100, 104-120, 136-152, 159-166 and 150-163 of Seq ID No 220; 15-27, 65-71, 77-99, 104-121, 128-154, 183-216, 223-229, 234-255, 277-287, 296-308 and 77-97 of Seq ID No 221; 8-18, 44-76, 102-109 and 49-57 of Seq ID No 222; 5-14, 28-40, 42-51, 54-60, 77-83, 89-100, 117-124, 146-172, 176-204, 216-231, 237-244, 267-278, 324-334, 342-348, 396-401, 427-433, 438-450, 452-457, 465-471, 473-481, 491-500, 509-515, 523-544, 550-556, 558-569, 589-595, 606-618, 625-632, 640-649, 665-671, 678-688, 691-698, 717-723, 728-734, 781-789, 800-805, 812-821, 833-868, 873-879, 889-905, 929-939, 988-998, 1046-1061, 1073-1079, 1089-1096, 1115-1124, 1132-1140, 1172-1196, 1220-1226, 1231-1249, 1269-1277, 1287-1301, 1307-1330, 1350-1361, 1369-1378, 1387-1412, 1414-1420, 1422-1439, 1484-1491, 1513-1529, 1552-1561, 1576-1583, 1606-1613, 1617-1640, 1647-1654, 1665-1679, 1686-1698, 1709-1727, 1736-1743, 1750-1757, 1771-1790, 1801-1807, 1817-1823, 1831-1842, 1859-1868, 1870-1882, 1884-1891, 1900-1906, 1909-1914, 1929-1935, 1952-1960, 1974-1988, 2002-2011, 2032-2063, 2071-2081, 2116-2124, 2139-2147, 2149-2159, 2163-2190, 2209-2215, 2245-2253, 2282-2287, 2331-2342, 2360-2370, 2379-2393, 2402-2408, 2414-2421, 2423-2430, 2433-2439, 2442-2450, 2472-2478, 2485-2493, 2495-2503, 2506-2512, 2547-2554, 2558-2564, 2615-2625, 2637-2652, 2692-2698, 2700-2706, 2711-2723, 2731-2740, 2748-2753, 2756-2762, 2765-2772, 2781-2798, 2810-2824, 2844-2852, 2885-2899, 2912-2922, 2937-2944, 2947-2970, 2988-2998, 3016-3025, 3032-3037, 3062-3071, 3129-3148, 3156-3161 and 530-607 of Seq ID No 223; 31-36, 57-62, 79-85, 90-96, 99-112, 120-146, 162-185, 193-203, 208-217, 219-226, 239-253, 283-290, 298-304, 306-321, 340-349, 351-361, 365-372, 386-395, 407-438, 473-486, 537-551, 558-568, 576-594, 598-604 and 75-95 of Seq ID No 224; 14-19, 24-30, 34-42, 45-52, 54-64, 66-82, 95-105, 107-118, 126-163, 171-177, 184-201, 210-215, 260-269, 273-279, 288-304, 321-327, 358-364, 370-375, 380-387, 394-404, 407-413, 421-431, 436-451, 465-474, 504-511, 531-552, 578-587, 614-626, 629-636, 638-671, 691-715, 719-729, 733-745, 752-759, 768-777, 785-792, 794-802, 805-824, 844-854, 867-880, 885-891, 893-902, 907-924, 939-948, 955-964, 966-975, 987-1000, 1012-1017, 1023-1028, 1050-1071, 1083-1098, 1102-1115, 1133-1146, 1170-1183, 1204-1211, 1213-1223, 1262-1311, 1313-1319, 1346-1355, 1366-1371, 1383-1405, 1409-1414 and 776-819 of Seq ID No 225; 12-27, 30-38, 54-61, 64-74, 82-96, 103-110, 117-125, 134-140, 147-158, 185-201, 218-225, 232-253, 265-280, 319-325, 350-362, 366-372, 376-386, 464-483, 485-490, 511-521, 531-537, 542-559, 564-574, 593-609, 613-619, 637-642, 668-677 and 195-214 of Seq ID No 226; 4-21, 59-67, 73-79, 84-91, 141-151, 186-197, 203-214, 222-227, 237-245, 255-260, 281-292, 294-311, 336-344, 346-355, 422-437, 459-466, 484-491 and 77-109 of Seq ID No 227; 10-45, 52-61, 63-70, 74-102, 112-122, 124-132, 164-178, 181-205, 212-240, 246-256 and 226-247 of Seq ID No 228; 38-50, 53-63, 78-87, 89-111, 126-152, 169-176, 179-186, 193-228, 254-267, 275-282, 288-304, 309-318, 325-341, 346-353, 358-367, 384-395, 404-427, 429-435, 456-465, 467-501, 510-521, 523-536,

541-548, 552-560, 563-584, 589-595, 597-620, 625-637, 639-645, 661-666, 712-729, 734-741, 743-750, 775-806, 809-816, 818-840, 842-850 and 693-714 of Seq ID No 229; 5-17, 30-37, 52-75, 77-86, 88-107, 112-135, 151-160, 178-222, 226-246, 263-270, 279-294, 306-314, 327-342, 345-352, 374-381, 389-416, 422-429, 435-449, 453-467, 473-500, 512-522, 524-531, 542-549, 552-560, 565-571, 575-586, 594-600, 613-619, 625-633, 635-641, 647-653, 667-674, 680-699, 711-729, 735-741, 764-775, 781-786, 792-798, 805-813, 817-825, 833-842, 850-855, 860-866, 869-910, 917-930, 949-990 and 533-562 of Seq ID No 230; 7-14, 39-46, 61-74, 83-89, 93-99, 110-121, 136-150, 172-180, 182-200, 207-216, 223-236, 238-251, 265-271, 280-288, 294-309, 320-336, 339-354, 362-377, 383-389, 401-407, 435-441, 446-453, 460-465, 472-487, 499-511, 518-528, 533-540, 557-570, 572-587, 631-637, 643-658, 663-669, 672-678, 681-687, 695-706, 714-728 and 118-139 of Seq ID No 231; 5-19, 24-30, 56-64, 69-79, 93-100, 102-111, 117-123, 125-133, 174-182, 185-199, 205-224, 268-275, 311-336 and 102-125 of Seq ID No 232; 6-35, 39-45, 57-62, 80-85, 92-106, 117-122, 126-171, 214-223, 253-260, 268-273, 285-291, 295-306, 315-320, 325-336, 361-366 and 172-202 of Seq ID No 233; 4-13, 24-37, 45-51, 58-66, 84-92, 112-121, 132-141, 151-171, 175-195, 204-212, 222-240, 262-268, 276-295, 305-336, 338-348, 354-362 and 160-183 of Seq ID No 234; 10-16, 24-35, 41-73, 78-104, 111-121, 124-139, 141-148, 150-164, 196-215, 224-241, 249-282, 299-307, 315-357, 368-378, 393-401 and 345-367 of Seq ID No 235; 4-32, 48-53, 61-67, 84-104, 112-118 and 106-130 of Seq ID No 236; 21-28, 31-36, 65-81, 98-105, 115-121, 123-131, 136-142, 155-161, 177-190 and 201-232 of Seq ID No 237; 4-15, 21-27, 33-39, 42-56, 58-64, 68-82, 84-90, 92-98, 113-122, 146-162, 168-175, 177-189, 191-203, 249-268, 279-285, 287-304, 328-342, 349-358, 371-378, 387-393, 404-413, 419-425, 467-479, 487-498, 513-524, 528-539, 541-565, 572-579, 595-606, 626-635, 637-642 and 612-626 of Seq ID No 238; 7-13, 52-70, 76-82, 97-106, 110-117 and 13-45 of Seq ID No 239; 5-10, 12-48, 59-64, 87-102, 107-128, 131-140, 154-161, 165-171, 173-215 and 54-74 of Seq ID No 240; 4-11, 19-28, 34-40, 74-81, 87-98, 126-147, 163-171, 184-193, 205-213 and 49-124 of Seq ID No 241; 7-14, 23-29, 35-40, 61-67, 99-106, 111-122, 124-133, 135-161, 187-206, 216-229, 236-245, 262-268, 271-280 and 256-273 of Seq ID No 242; 4-13, 17-37, 47-54, 85-99, 105-113, 120-132, 147-166, 180-186, 192-199, 204-216 and 127-144 of Seq ID No 243; 14-27, 29-37, 52-62, 68-76, 89-96, 117-123, 125-131, 137-145, 166-195, 205-212, 214-222, 228-235, 258-264, 271-281, 288-296, 308-324, 332-339, 355-361, 365-371 and 268-293 of Seq ID No 244; 4-21, 30-42, 54-60, 78-85, 90-110, 141-147, 160-168, 176-185, 194-206, 218-225, 230-245, 251-261, 287-293, 295-304, 320-326, 334-347, 351-362, 386-402, 413-423, 427-433, 439-453, 456-477, 480-493, 507-513, 526-539, 574-581, 591-598, 600-609, 614-632, 655-665, 685-691, 703-712, 742-747, 757-775, 797-803, 813-819, 823-829, 880-887, 901-906, 930-944, 948-958, 962-968, 971-995, 1002-1009, 1017-1023, 1036-1053, 1069-1081, 1107-1124, 1129-1152, 1178-1195, 1211-1223, 1249-1266, 1271-1288, 1334-1340, 1346-1367, 1-63 and 171-189 of Seq ID No 245; 4-22, 52-63, 70-75, 94-104, 112-125, 133-141, 176-199, 209-216, 244-259, 287-299, 336-352, 366-372, 386-399, 421-436, 444-449, 457-466, 481-487, 506-529, 531-540 and 295-378 of Seq ID No 246; 9-30, 43-49, 58-75, 86-96, 119-131, 138-147, 162-167, 181-201, 208-214 and 16-121 of Seq ID No 247; 4-27, 52-58, 80-90, 92-100, 108-114, 118-143, 169-176, 189-198, 247-261, 281-287, 307-317, 323-329, 352-363, 372-381, 396-411, 413-426, 429-440, 442-450, 456-461, 468-479 and 1-73 of Seq ID No 248; 4-32, 47-52, 57-63, 71-78, 92-104, 126-142, 153-175 and 145-163 of Seq ID No 249; 17-23, 35-41, 51-70, 73-86, 104-125 and 105-129 of Seq ID No 250; 25-32, 41-50, 75-85, 87-103, 115-122, 138-149, 164-171, 188-210, 212-220, 224-234, 256-273, 288-299, 304-310, 330-336, 357-365, 382-390, 399-405, 414-421, 440-446, 454-461, 480-486, 502-514, 518-540, 543-553, 561-567, 572-580, 582-588, 595-630, 633-651, 672-681, 691-709, 760-767, 813-832, 841-848, 852-866, 873-893, 919-925, 927-933, 940-955, 957-978, 984-997, 1000-1010, 1035-1040, 1044-1051, 1058-1064, 1081-1091, 1097-1124, 1129-1138, 1144-1150, 1158-1165, 1170-1180, 909-936 and 1001-1031 of Seq ID No 251; 4-12, 19-26, 31-41, 49-64, 66-86, 101-117, 119-127, 134-142, 152-161, 163-172, 179-188, 209-218, 234-241, 276-291, 294-300, 307-320, 324-341, 346-356, 373-387, 389-397, 410-416, 418-436, 444-454, 460-472, 481-486, 500-507, 511-535, 541-549, 553-559, 579-586, 602-607, 613-620, 628-640, 654-663, 671-678, 681-691, 709-722, 741-754, 766-774, 778-786, 797-803 and 212-226 of Seq ID No 252; 4-10, 15-27, 34-54, 60-73, 79-88, 101-115, 120-136, 154-162, 167-172, 222-240 and 126-195 of Seq ID No 253; 5-16, 18-25, 29-35, 57-63, 86-91, 107-121, 123-131, 170-179, 185-199, 204-226, 250-255, 262-274, 291-296, 325-347 and 1-38 of Seq ID No 254; 7-19, 22-34, 36-42, 48-54, 60-66, 71-76, 104-110, 118-133, 135-145, 158-164, 167-174, 182-193, 196-204, 217-229, 251-290, 293-299, 309-315 and 288-318 of Seq ID No 255; 43-51, 55-61, 66-

73, 80-90, 103-127, 133-142, 174-180, 185-196, 203-210, 229-235, 239-251, 258-266, 272-278, 289-314, 316-326, 340-346, 355-361 and 14-27 of Seq ID No 256; 4-25, 27-33, 35-41, 52-74, 76-89, 99-124, 138-144, 146-159, 167-182, 184-191, 193-206, 211-223, 232-240, 249-257, 270-279, 281-287, 293-310, 322-341, 347-356 and 292-322 of Seq ID No 257; 5-13, 28-38, 43-60, 67-72, 98-116, 122-134, 137-151, 167-174, 177-195, 197-216 and 99-195 of Seq ID No 258; 15-33, 35-42, 48-57, 62-68, 73-91, 107-119, 121-153, 173-194, 205-210, 223-228, 234-241, 243-259, 275-298, 308-315, 327-340, 342-370, 376-391, 398-404, 410-419 and 71-122 of Seq ID No 259; 12-39, 43-64, 87-95, 99-105, 114-126, 128-136, 139-147, 212-225 and 107-141 of Seq ID No 260; 6-33, 40-45, 60-75, 79-86, 121-129, 131-137, 161-167, 172-178, 186-195, 203-212, 236-244, 257-264, 278-294, 306-312, 345-358, 368-381, 386-395, 404-410, 412-418 and 198-270 of Seq ID No 261; 18-31, 34-41, 50-56, 69-83, 99-106, 129-141, 147-153, 159-168, 170-178, 190-198, 200-212, 221-232, 237-255, 261-266, 274-292 and 118-216 of Seq ID No 262; 17-47, 61-67, 87-93, 115-121, 126-132, 140-148, 167-173, 179-186, 214-223, 250-255, 264-272, 282-294, 306-318, 338-353, 358-377, 385-401, 414-420, 433-441, 451-457, 470-480, 505-511, 544-550, 571-581, 600-607, 612-618, 631-648, 655-662, 669-681, 693-714, 721-726, 733-740, 757-778, 813-823, 831-838, 851-857, 866-876, 893-905, 912-917, 930-936, 951-963, 971-981, 1008-1019, 1021-1033, 1035-1041, 1054-1064, 1066-1076, 1097-1110, 1113-1121, 1126-1140, 1159-1171, 1182-1195, 1197-1203, 1216-1222, 1231-1240, 1243-1262, 1268-1287 and 738-828 of Seq ID No 263; 19-28, 40-46, 51-57, 68-74, 81-87, 98-108, 111-121 and 20-36 of Seq ID No 264; 4-17, 19-44, 60-69, 80-86, 110-116 and 33-60 of Seq ID No 265; 8-16, 18-28, 42-50, 53-75, 79-86, 94-99, 122-128, 136-142, 149-163, 166-173, 198-212, 254-272, 288-295, 304-318, 324-329, 343-348, 351-364, 367-383, 389-395, 411-420, 427-436 and 11-56 of Seq ID No 266; 19-25 and 6-24 of Seq ID No 267; 6-39, 59-68 and 44-63 of Seq ID No 268; 5-14, 21-28, 38-53 and 29-41 of Seq ID No 269; 4-13, 31-41, 56-65 and 32-56 of Seq ID No 270; 5-12 and 4-21 of Seq ID No 271; 4-18 and 17-32 of Seq ID No 272; 4-10, 23-33 and 14-30 of Seq ID No 273; 26-34, 44-53 and 35-52 of Seq ID No 274; 1-19 of Seq ID No 275; 4-17, 23-30, 32-37 and 6-23 of Seq ID No 276; 5-33, 40-58, 61-66 and 45-66 of Seq ID No 277; 15-41, 61-67 and 41-65 of Seq ID No 278; 4-12, 16-23, 26-37 and 10-29 of Seq ID No 279; 23-39 and 37-55 of Seq ID No 280; 12-20 and 38-55 of Seq ID No 281; 22-37 and 7-22 of Seq ID No 282; 3-14 of Seq ID No 283; 6-16, 43-65, 71-76 and 17-31 of Seq ID No 284; 4-13, 27-39, 42-69 and 17-32 of Seq ID No 285; 4-12, 26-39 and 10-25 of Seq ID No 286; 2-31 of Seq ID No 287; 6-38, 49-62 and 39-55 of Seq ID No 288; 4-10, 24-30 and 2-19 of Seq ID No 289; 12-17, 25-46 and 15-30 of Seq ID No 290; 4-13 and 2-28 of Seq ID No 291; 30-38 and 17-45 of Seq ID No 292 ; 24-33, 55-61 and 31-61 of Seq ID No 293; 4-26, 34-48 and 15-33 of Seq ID No 294; 9-15 and 1-22 of Seq ID No 295; 4-31 and 14-33 of Seq ID No 296; 5-34, 49-55, 64-82 and 69-83 of Seq ID No 297; 33-45 and 21-39 of Seq ID No 298; 7-14, 24-32, 42-65, 79-86 and 50-64 of Seq ID No 299; 13-27, 33-43, 45-62 and 12-37 of Seq ID No 300; 4-15, 17-32 and 10-26 of Seq ID No 301; 4-9, 11-43, 45-75 and 47-69 of Seq ID No 302; 4-18, 22-37 and 17-34 of Seq ID No 303; 4-14 and 5-24 of Seq ID No 304; 7-33, 35-46 and 1-19 of Seq ID No 305; 13-37, 69-75 and 51-69 of Seq ID No 306; 14-24, 26-34, 37-49, 66-78 and 2-25 of Seq ID No 307; 17-46, 52-57, 59-64 and 54-68 of Seq ID No 308; 4-22 and 13-25 of Seq ID No 309; 8-40, 53-63 and 29-50 of Seq ID No 310; 16-28 and 32-40 of Seq ID No 311; 14-20, 22-28, 39-45 and 2-22 of Seq ID No 312; 4-13 and 12-31 of Seq ID No 313; 15-21 and 2-17 of Seq ID No 314; 4-17 and 20-36 of Seq ID No 315; 4-19 and 9-18 of Seq ID No 316; 4-14 and 3-19 of Seq ID No 317; 4-21, 32-40 and 21-39 of Seq ID No 318; 4-13 and 10-27 of Seq ID No 319; 18-31, 39-47, 75-87, 89-98 and 79-99 of Seq ID No 320; 15-21 and 9-24 of Seq ID No 321; 4-14, 18-27, 30-53, 55-64, 68-74, 81-98 and 22-40 of Seq ID No 322; 7-24, 44-51 and 35-60 of Seq ID No 323; 10-47 and 23-37 of Seq ID No 324; 4-10, 12-46 and 7-22 of Seq ID No 325; 20-27 and 1-13 of Seq ID No 326; 6-19, 41-51 and 9-37 of Seq ID No 327; 4-9, 11-17 and 9-23 of Seq ID No 328; 4-17, 23-38, 46-66, 68-85 and 34-46 of Seq ID No 329; 4-18, 34-59, 75-81 and 61-84 of Seq ID No 330; 6-17 and 7-28 of Seq ID No 331; 4-32, 56-61 and 35-52 of Seq ID No 332; 4-14, 27-71, 74-88, 93-110, 115-120, 124-130, 139-154, 161-172 and 146-171 of Seq ID No 333; 4-21 and 3-15 of Seq ID No 334; 12-17 and 9-26 of Seq ID No 335; 10-21, 45-58 and 51-67 of Seq ID No 336; 59-66, 68-84 and 13-42 of Seq ID No 337; 11-16 and 1-16 of Seq ID No 338; 4-19, 23-37 and 10-30 of Seq ID No 339; 19-27, 35-46, 48-66, 82-88, 99-105, 113-119 and 42-59 of Seq ID No 340; 135-147 of Seq ID No 171; 658-682 of Seq ID No 172; 411-427 and 1226-1246 of Seq ID No 173; 794-817 and 801-824 of Seq ID No 174; 468-492 and 474-495 of Seq ID No 177; 366-388 of Seq

ID No 178; 266-291, 287-312 and 308-333 of Seq ID No 179; 197-213 and 195-211 of Seq ID No 180; 252-275, 262-285 and 812-830 of Seq ID No 181; 94-112, 97-120 and 104-128 of Seq ID No 182; 257-281 of Seq ID No 183; 106-134 of Seq ID No 184; 70-86 of Seq ID No 187; 358-383, 378-402, 397-421, 499-524, 520-545, 541-566, 622-646, 641-665 and 660-684 of Seq ID No 189; 248-260 of Seq ID No 192; 15-34 of Seq ID No 194; 112-129 of Seq ID No 195; 333-358 and 353-378 of Seq ID No 196; 316-343, 339-366 and 362-389 of Seq ID No 197; 98-123 and 104-126 of Seq ID No 198; 20-43 and 23-48 of Seq ID No 201; 124-145 of Seq ID No 203; 717-738 of Seq ID No 204; 37-56 of Seq ID No 205; 118-134 of Seq ID No 206; 500-522 of Seq ID No 208; 32-47 of Seq ID No 211; 25-51, 47-73 and 69-95 of Seq ID No 212; 503-529 of Seq ID No 215; 112-128 of Seq ID No 216; 181-199 of Seq ID No 218; 109-121 of Seq ID No 219; 150-163 of Seq ID No 220; 77-97 of Seq ID No 221; 564-586 of Seq ID No 223; 75-94 of Seq ID No 224; 776-798, 784-808 and 794-815 of Seq ID No 225; 196-212, 78-100 and 85-107 of Seq ID No 226; 536-553 of Seq ID No 230; 102-125 of Seq ID No 232; 178-198 of Seq ID No 233; 612-626 of Seq ID No 238; 171-187 of Seq ID No 245; 296-320, 315-339, 334-358 and 353-377 of Seq ID No 246; 47-71 of Seq ID No 247; 1-25, 20-45 and 40-64 of Seq ID No 248; 146-161 of Seq ID No 249; 910-935 and 1007-1030 of Seq ID No 251; 212-226 of Seq ID No 252; 126-152, 148-173 and 169-195 of Seq ID No 253; 288-310 and 293-316 of Seq ID No 255; 293-312 of Seq ID No 257; 154-170 of Seq ID No 258; 72-95, 90-112 and 97-121 of Seq ID No 259; 135-150 and 146-163 of Seq ID No 262; 799-827 of Seq ID No 263; 23-43 and 33-53 of Seq ID No 266; 44-62 of Seq ID No 268; 6-22 of Seq ID No 276; 37-54 of Seq ID No 280; 40-54 of Seq ID No 281; 7-21 of Seq ID No 282; 4-11, 16-34, 48-55, 67-77, 87-106 and 153-183 of Seq ID No 425; 22-40, 49-65, 70-91, 95-109, 111-125, 146-207, 209-216, 219-225, 229-244, 251-270, 274-286, 292-309, 316-329, 335-355, 358-370, 376-388, 392-419, 425-430, 435-441, 448-455, 464-478, 486-515 and 437-465 of Seq ID No 426; 5-19, 25-31, 43-48, 60-79, 88-100, 105-129, 148-171, 187-193, 243-263, 316-322, 334-340, 345-351, 369-378, 381-391, 399-404, 474-483, 502-517, 525-530, 558-568, 579-596, 622-627, 631-638, 644-651, 653-660, 674-680, 687-693, 721-728, 743-753, 760-775, 788-795, 806-813, 821-828, 835-842, 847-859, 868-887 and 300-347 of Seq ID No 427; 5-26, 37-44, 89-97, 112-118, 121-128, 138-154, 157-165, 176-181, 188-198, 205-218, 223-243, 247-253, 260-279 and 76-155 of Seq ID No 428; 4-29, 41-46, 49-68, 82-88, 121-147, 158-164, 187-193, 195-208, 229-236, 244-249, 251-263, 269-275, 307-313, 337-343, 348-381, 392-398, 402-408, 432-438, 85-117 and 194-239 of Seq ID No 429; 5-12, 14-22, 28-34, 40-46, 70-79, 84-129, 152-165, 174-182 and 37-109 of Seq ID No 430; 5-16, 18-52, 54-72, 81-86, 118-126, 136-145, 151-157, 168-180, 209-233, 244-270, 295-302, 315-326, 329-337, 345-352, 364-373, 397-402, 408-418, 424-431, 435-443, 472-480, 483-489, 504-510, 519-527, 549-564, 576-599, 605-637, 641-673 and 91-98 of Seq ID No 431; 23-36, 42-52, 133-140, 151-157, 242-247, 267-277, 295-301, 320-328, 333-339, 345-352, 365-371, 397-403, 415-428, 456-465, 481-487, 489-495, 508-516, 518-527, 585-592, 606-614, 631-637, 643-658, 665-670, 723-728, 737-744, 752-759, 787-793, 835-841, 873-885, 918-928, 938-945, 951-966, 978-988, 1015-1020, 1030-1036, 1044-1052, 1058-1069, 1071-1079, 1081-1088, 1113-1119, 1125-1138, 1141-1147, 1164-1170, 1172-1177, 1190-1200, 1214-1220, 1230-1236, 1239-1245, 1262-1268, 1270-1275, 1288-1298, 1312-1318, 1328-1334, 1337-1343, 1360-1366, 1368-1373, 1386-1396, 1410-1416, 1426-1432, 1435-1441, 1458-1464, 1466-1471, 1484-1494, 1508-1514, 1524-1530, 1533-1539, 1556-1562 and 307-340 of Seq ID No 432; 19-25, 35-41, 44-50, 66-72, 74-79, 92-102, 116-122, 132-138, 141-147, 164-170, 172-177, 190-200, 214-220, 230-236, 239-245, 262-268, 270-275, 288-298, 312-318, 328-334, 337-343, 360-366, 368-373, 386-396, 410-416, 426-432, 435-441, 458-464, 466-478, 504-524, 79-148, 177-246, 275-344 and 373-442 of Seq ID No 433; 7-14, 16-23, 33-39, 46-53, 72-79, 92-115, 123-130, 156-175, 179-187, 214-220, 239-246, 266-274, 302-325, 338-354, 360-370, 375-390, 392-401, 421-428, 430-463 and 29-58 of Seq ID No 434; 4-9, 22-39, 58-65, 72-82, 87-92, 99-104, 107-119, 143-166, 171-177, 194-202, 205-213, 220-228, 231-240, 247-263, 309-315, 317-323, 336-343 and 294-320 of Seq ID No 435; 4-10, 12-18, 24-29, 34-43, 50-65, 70-76, 111-117, 129-138, 152-159, 166-171, 184-195, 200-210, 224-236, 241-251, 274-283, 285-296, 313-319, 332-341, 348-355, 378-386, 410-416, 433-445, 475-482, 523-529, 531-540, 584-596, 626-633, 674-680, 682-688, 738-750, 780-787, 828-834, 836-842, 853-862, 882-887, 893-912 and 604-676 of Seq ID No 436; 15-38, 49-57, 60-99, 103-119, 124-194, 200-206, 215-249, 251-291, 307-313, 315-347, 369-378, 383-390, 393-400, 405-411, 423-435, 440-447, 454-460, 470-486, 490-503, 532-539, 542-549, 551-567, 579-592 and 509-583 of Seq ID No 437; 38-44, 47-88, 95-103, 157-172, 235-240, 250-260, 263-276, 294-300, 312-317, 331-337, 369-

391, 412-419, 442-448, 453-463, 490-529, 537-555, 571-580, 600-617, 619-627, 642-648, 682-687, 693-700, 716-722, 738-748, 756-763, 779-789, 796-802, 820-828, 833-840, 846-853, 862-872, 880-887, 894-899, 924-937, 957-963, 1006-1012, 1043-1049, 1063-1069, 1076-1097 and 124-147 of Seq ID No 438; 4-28, 31-49, 60-71, 75-102, 104-114, 134-144, 160-184, 250-257, 277-285, 287-294, 330-338, 345-351, 367-374, 381-388, 393-399, 402-407, 420-426, 443-448, 458-464, 411-436 and 454-488 of Seq ID No 439; 20-27, 45-55, 57-64, 66-77, 98-106, 130-137, 155-165, 167-174, 176-187, 194-203, 208-223, 227-238, 245-251, 257-270, 273-278, 287-299, 330-345, 352-358, 363-385, 392-399, 410-417, 437-443, 467-484, 486-492, 495-500, 504-516, 526-536 and 219-270 of Seq ID No 440; 11-22, 24-31, 46-63, 65-71, 73-88, 95-109, 174-181, 183-201, 204-212, 216-222, 228-233, 241-247 and 142-221 of Seq ID No 441; 8-28, 51-59, 67-84, 93-98, 140-152, 154-162, 183-188 and 91-125 of Seq ID No 442; 10-22, 27-61 and 69-100 of Seq ID No 443; 7-15, 18-26, 94-100, 126-131, 152-165, 219-228, 254-263, 274-292, 297-308, 333-340, 342-352, 354-371, 373-379, 403-410, 420-438, 450-456, 463-470, 489-495, 503-512 and 97-173 of Seq ID No 444; 4-21, 37-43, 49-65, 67-74, 76-90, 113-119, 131-141, 155-173, 175-189, 192-199, 207-221, 247-254, 266-276, 317-322, 337-343, 387-393, 408-428, 439-448, 451-460, 469-479, 482-487, 493-501, 517-523, 533-542 and 480-503 of Seq ID No 445; 11-26, 40-46, 78-86, 93-103, 121-126, 132-138, 166-177, 183-196, 203-212, 214-221, 228-263, 304-311, 323-338, 345-351, 357-363, 379-393, 420-434, 442-448, 518-527, 547-553, 581-591, 602-609, 637-645, 665-674, 687-692, 701-708, 730-739, 796-802, 844-857, 882-888, 903-914, 944-950, 976-983, 1027-1033, 1049-1057, 1066-1072, 1085-1092, 1120-1127, 1137-1144, 1153-1158, 1165-1176, 1181-1187, 1221-1230, 1238-1244, 1269-1274 and 605-632 of Seq ID No 446; 6-47, 57-65, 83-95, 109-121, 138-147, 154-164, 167-177, 194-200, 202-212, 227-234, 240-253, 260-267, 283-291, 320-329, 340-347, 356-364, 412-422, 430-436, 441-459, 465-475, 478-486, 498-507 and 59-84 of Seq ID No 447; 10-21, 58-83, 88-97, 120-126 and 21-51 of Seq ID No 448; 5-39, 56-62, 76-88, 90-114, 138-162, 170-195, 202-221, 228-250, 264-270, 304-355, 374-387, 391-416, 462-471, 526-546, 554-561, 574-579, 639-645, 651-660, 674-682, 689-694 and 666-697 of Seq ID No 449; 6-30, 36-42, 143-157, 176-197, 202-209, 216-233, 241-246, 275-287, 292-299, 315-325, 343-350, 375-380, 397-403, 411-420, 422-434, 441-448, 467-474, 477-499, 555-568, 591-597, 601-609, 623-644, 667-688, 692-698, 703-718, 736-747, 757-766, 782-791, 795-801, 832-840, 859-865 and 226-269 of Seq ID No 450; 6-23, 43-51, 61-67, 73-82, 91-97, 123-130, 149-158, 164-175, 228-234, 240-246, 248-255, 262-272, 326-332, 340-347, 365-371, 377-388, 409-419, 425-431, 438-445, 449-457, 464-470, 496-507, 559-568, 575-581, 603-608, 617-623, 633-639, 648-654, 659-670, 695-701, 734-752, 806-814, 816-829, 861-868, 891-899, 904-909, 937-945, 947-960, 978-983, 992-999, 1022-1031, 1068-1076, 1078-1091, 1109-1114, 1123-1130, 1153-1162, 1199-1207, 1209-1222, 1254-1261, 1284-1293, 1330-1338, 1340-1353, 1371-1376, 1385-1392, 1415-1421, 1433-1438, 1460-1465, 1470-1492 and 1422-1458 of Seq ID No 451; 82-94, 111-118, 125-131, 206-212, 261-266, 310-320, 328-338, 345-351, 353-360, 414-420, 424-434, 440-447, 451-500, 506-516, 548-561, 566-572, 584-591, 601-622, 630-636, 650-659, 661-667, 674-699, 703-711, 717-729, 736-744, 752-759, 765-771, 813-822, 826-842, 852-868, 870-877, 881-895, 897-906, 913-922 and 602-671 of Seq ID No 452; 12-18, 20-25, 43-54, 56-65, 73-79, 82-88, 99-111, 136-142, 153-169, 171-183, 195-223, 229-248, 255-260, 272-277, 281-292, 298-319, 322-329, 332-351, 363-379, 381-389 and 275-304 of Seq ID No 453; 4-9, 34-48, 65-77, 101-106, 111-131, 138-153, 186-191, 230-250 and 148-219 of Seq ID No 454; 4-23, 30-35, 42-53, 67-76, 82-87, 101-108, 112-130, 132-138, 147-152, 161-183, 187-208, 218-225, 265-283, 295-303, 306-317, 322-334, 338-357, 360-368, 370-383, 387-398, 400-419, 421-430, 104-182 and 240-304 of Seq ID No 455; 4-12, 63-69, 94-102, 146-164, 166-173, 175-181, 193-207, 263-281, 286-295, 301-306, 330-343, 369-378, 382-388, 414-420, 422-430, 438-454, 456-462, 472-531, 543-560, 581-591, 596-605, 614-623, 626-635, 656-662, 669-676, 683-690, 693-698, 705-711, 728-736, 752-764 and 69-102 of Seq ID No 456; 6-12, 43-53, 141-147, 164-179, 185-195, 197-206, 227-235, 237-271, 288-305, 308-317, 335-341, 351-357, 365-376, 386-395, 397-416, 422-447 and 11-35 of Seq ID No 457; 16-24, 50-65, 73-84, 88-99, 114-124, 130-146, 181-187, 193-203, 214-220, 236-247, 250-258, 287-297 and 50-113 of Seq ID No 458; 4-25, 50-55, 76-82, 117-123, 131-137, 139-148, 157-166, 239-245, 253-258, 266-275, 277-292, 300-306, 51-83 and 93-161 of Seq ID No 459; 6-22, 34-43, 51-86, 93-100, 110-116, 150-161, 164-171, 180-187, 197-218 and 168-237 of Seq ID No 460; 4-27, 55-60, 74-82 and 10-46 of Seq ID No 461; 6-19, 25-31, 43-49, 60-79, 88-100, 105-129, 148-161, 164-171, 187-193, 243-263, 316-322, 334-340, 369-378, 381-391, 398-404, 460-466, 474-483, 502-509, 511-517, 525-530, 558-567, 579-596, 622-627, 631-638, 641-651, 653-659, 674-680, 687-693, 710-716,

720-727, 743-753, 760-775, 788-795, 806-813, 821-828, 836-842, 847-860, 865-880 and 258-377 of Seq ID No 462; 4-11, 25-64, 71-79, 88-94, 107-120, 123-132, 167-188, 231-237, 240-246, 261-267, 306-311, 330-342, 351-358, 389-395, 406-418, 429-434, 439-448, 483-501, 511-520 and 71-143 of Seq ID No 463; 4-18, 22-27, 53-64, 94-100, 121-127, 133-139, 155-164, 177-182, 187-196, 206-218, 224-242, 248-253, 258-277 and 184-253 of Seq ID No 464; 10-17, 56-67, 72-82, 94-99, 106-113, 166-173, 229-235, 243-283, 295-301, 313-321, 326-331, 342-348, 396-414, 423-435, 446-452, 454-462, 496-502, 511-534, 543-556, 563-570, 586-593, 616-626, 638-645, 653-662, 679-696, 731-737, 766-774, 776-782, 790-796, 810-817, 825-835, 837-846 and 540-615 of Seq ID No 465; 13-24, 30-36, 73-81, 89-95, 109-115, 117-143, 161-173, 179-189, 226-244, 251-261, 275-281, 298-305, 307-315, 323-328, 364-374, 69-186 and 264-354 of Seq ID No 466; 19-25 and 6-22 of Seq ID No 467; 6-39, 59-68 and 43-62 of Seq ID No 468; 6-14, 22-32 and 1-27 of Seq ID No 469; 4-41 and 28-40 of Seq ID No 470; 8-14 and 4-19 of Seq ID No 471; 4-10, 12-22, 30-35 and 6-33 of Seq ID No 472; 4-16, 24-33 and 37-54 of Seq ID No 473; 2-23 of Seq ID No 474; 4-21, 27-33, 36-41 and 14-34 of Seq ID No 475; 4-14, 24-30, 37-42, 57-78, 83-89, 94-103, 113-131 and 100-122 of Seq ID No 476.

15. A process for producing a E. faecalis hyperimmune serum reactive antigen or a fragment thereof according to any one of the claims 11 to 14 comprising expressing the nucleic acid molecule according to any one of claims 1 to 7.

16. A process for producing a cell, which expresses a E. faecalis hyperimmune serum reactive antigen or a fragment thereof according to any one of the claims 11 to 14 comprising transforming or transfecting a suitable host cell with the vector according to claim 8 or claim 9.

17. A pharmaceutical composition, especially a vaccine, comprising a hyperimmune serum-reactive antigen or a fragment thereof, as defined in any one of claims 11 to 14 or a nucleic acid molecule according to any one of claims 1 to 7.

18. A pharmaceutical composition, especially a vaccine, according to claim 17, characterized in that it further comprises an immunostimulatory substance, preferably selected from the group comprising polycationic polymers, especially polycationic peptides, immunostimulatory deoxynucleotides (ODNs), peptides containing at least two LysLeuLys motifs, neuroactive compounds, especially human growth hormone, alumn, Freund's complete or incomplete adjuvants or combinations thereof.

19. Use of a nucleic acid molecule according to any one of claims 1 to 7 or a hyperimmune serum-reactive antigen or fragment thereof according to any one of claims 11 to 14 for the manufacture of a pharmaceutical preparation, especially for the manufacture of a vaccine against enterococcal infection.

20. An antibody, or at least an effective part thereof, which binds at least to a selective part of the hyperimmune serum-reactive antigen or a fragment thereof according to any one of claims 11 to 14.

21. An antibody according to claim 20, wherein the antibody is a monoclonal antibody.

22. An antibody according to claim 20 or 21, wherein said effective part comprises Fab fragments.

23. An antibody according to any one of claims 20 to 22, wherein the antibody is a chimeric antibody.

24. An antibody according to any one of claims 20 to 23, wherein the antibody is a humanized antibody.

25.   A hybridoma cell line, which produces an antibody according to any one of claims 20 to 24.

26.   A method for producing an antibody according to claim 20, characterized by the following steps:
   • initiating an immune response in a non-human animal by administrating an hyperimmune serum-reactive antigen or a fragment thereof, as defined in any one of the claims 11 to 14, to said animal,
   • removing an antibody containing body fluid from said animal, and
   • producing the antibody by subjecting said antibody containing body fluid to further purification steps.

27.   Method for producing an antibody according to claim 21, characterized by the following steps:
   • initiating an immune response in a non-human animal by administrating an hyperimmune serum-reactive antigen or a fragment thereof, as defined in any one of the claims 12 to 15, to said animal,
   • removing the spleen or spleen cells from said animal,
   • producing hybridoma cells of said spleen or spleen cells,
   • selecting and cloning hybridoma cells specific for said hyperimmune serum-reactive antigens or a fragment thereof,
   • producing the antibody by cultivation of said cloned hybridoma cells and optionally further purification steps.

28.   Use of the antibodies according to any one of claims 20 to 24 for the preparation of a medicament for treating or preventing enterococcal infections.

29.   An antagonist which binds to the hyperimmune serum-reactive antigen or a fragment thereof according to any one of claims 11 to 14.

30.   A method for identifying an antagonist capable of binding to the hyperimmune serum-reactive antigen or fragment thereof according to any one of claims 11 to 14 comprising:
   a) contacting an isolated or immobilized hyperimmune serum-reactive antigen or a fragment thereof according to any one of claims 11 to 14 with a candidate antagonist under conditions to permit binding of said candidate antagonist to said hyperimmune serum-reactive antigen or fragment, in the presence of a component capable of providing a detectable signal in response to the binding of the candidate antagonist to said hyperimmune serum reactive antigen or fragment thereof; and
   b) detecting the presence or absence of a signal generated in response to the binding of the antagonist to the hyperimmune serum reactive antigen or the fragment thereof.

31.   A method for identifying an antagonist capable of reducing or inhibiting the interaction activity of a hyperimmune serum-reactive antigen or a fragment thereof according to any one of claims 11 to 14 to its interaction partner comprising:
   a) providing a hyperimmune serum reactive antigen or a hyperimmune fragment thereof according to any one of claims 11-14,
   b) providing an interaction partner to said hyperimmune serum reactive antigen or a fragment thereof, especially an antibody according to any one of the claims 20 to 24,
   c) allowing interaction of said hyperimmune serum reactive antigen or fragment thereof to said interaction partner to form a interaction complex,
   d) providing a candidate antagonist,
   e) allowing a competition reaction to occur between the candidate antagonist and the interaction complex,
   f) determining whether the candidate antagonist inhibits or reduces the interaction activities of the hyperimmune serum reactive antigen or the fragment thereof with the interaction partner.

32. Use of any of the hyperimmune serum reactive antigen or fragment thereof according to any one of claims 11 to 14 for the isolation and/or purification and/or identification of an interaction partner of said hyperimmune serum reactive antigen or fragment thereof.

33. A process for *in vitro* diagnosing a disease related to expression of the hyperimmune serum-reactive antigen or a fragment thereof according to any one of claims 11 to 14 comprising determining the presence of a nucleic acid sequence encoding said hyperimmune serum reactive antigen and fragment according to any one of claims 1 to 7 or the presence of the hyperimmune serum reactive antigen or fragment thereof according to any one of claims 11-14.

34. A process for *in vitro* diagnosis of a bacterial infection, especially a enterococcal infection, comprising analysing for the presence of a nucleic acid sequence encoding said hyperimmune serum reactive antigen and fragment according to any one of claims 1 to 7 or the presence of the hyperimmune serum reactive antigen or fragment thereof according to any one of claims 11 to 14.

35. Use of the hyperimmune serum reactive antigen or fragment thereof according to any one of claims 11 to 14 for the generation of a peptide binding to said hyperimmune serum reactive antigen or fragment thereof, wherein the peptide is selected from the group comprising anticalines.

36. Use of the hyperimmune serum-reactive antigen or fragment thereof according to any one of claims 11 to 14 for the manufacture of a functional nucleic acid, wherein the functional nucleic acid is selected from the group comprising aptamers and spiegelmers.

37. Use of a nucleic acid molecule according to any one of claims 11 to 14 for the manufacture of a functional ribonucleic acid, wherein the functional ribonucleic acid is selected from the group comprising ribozymes, antisense nucleic acids and siRNA.

Fig. 1

2/5
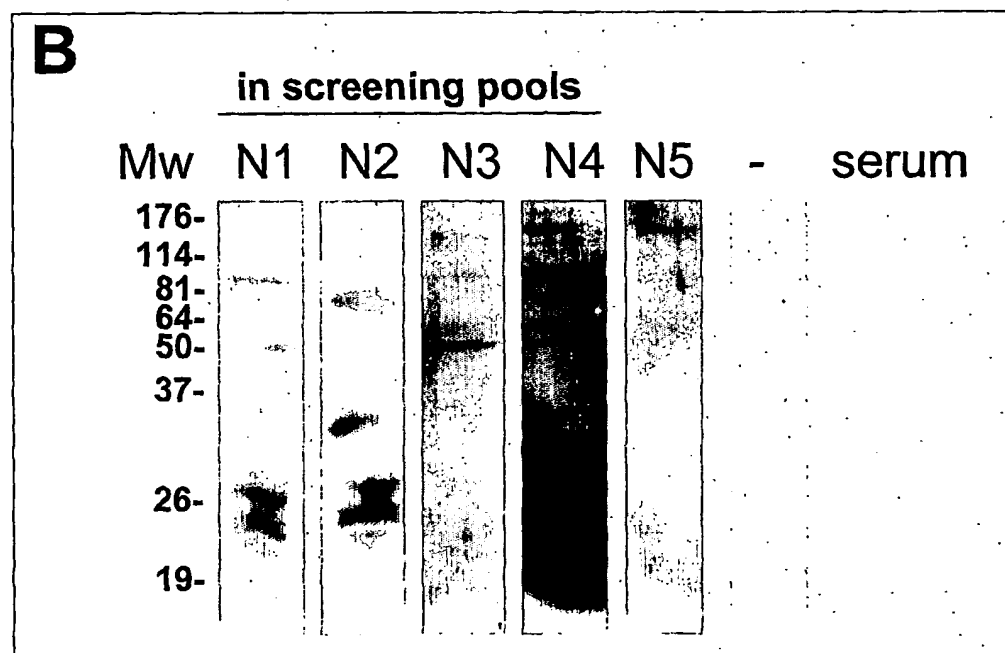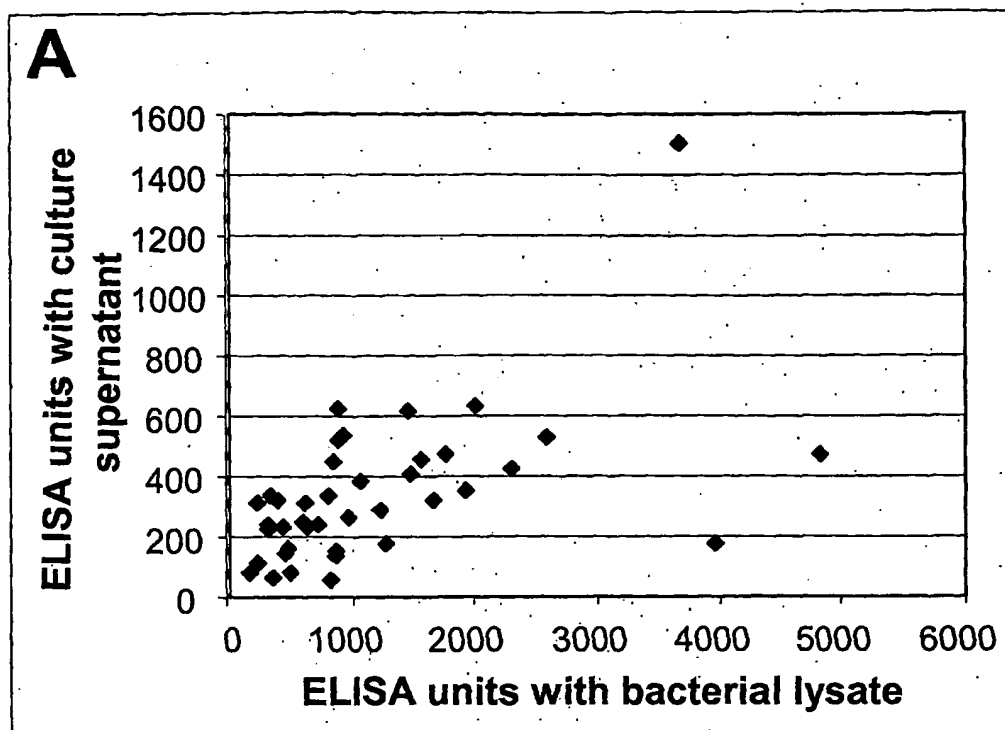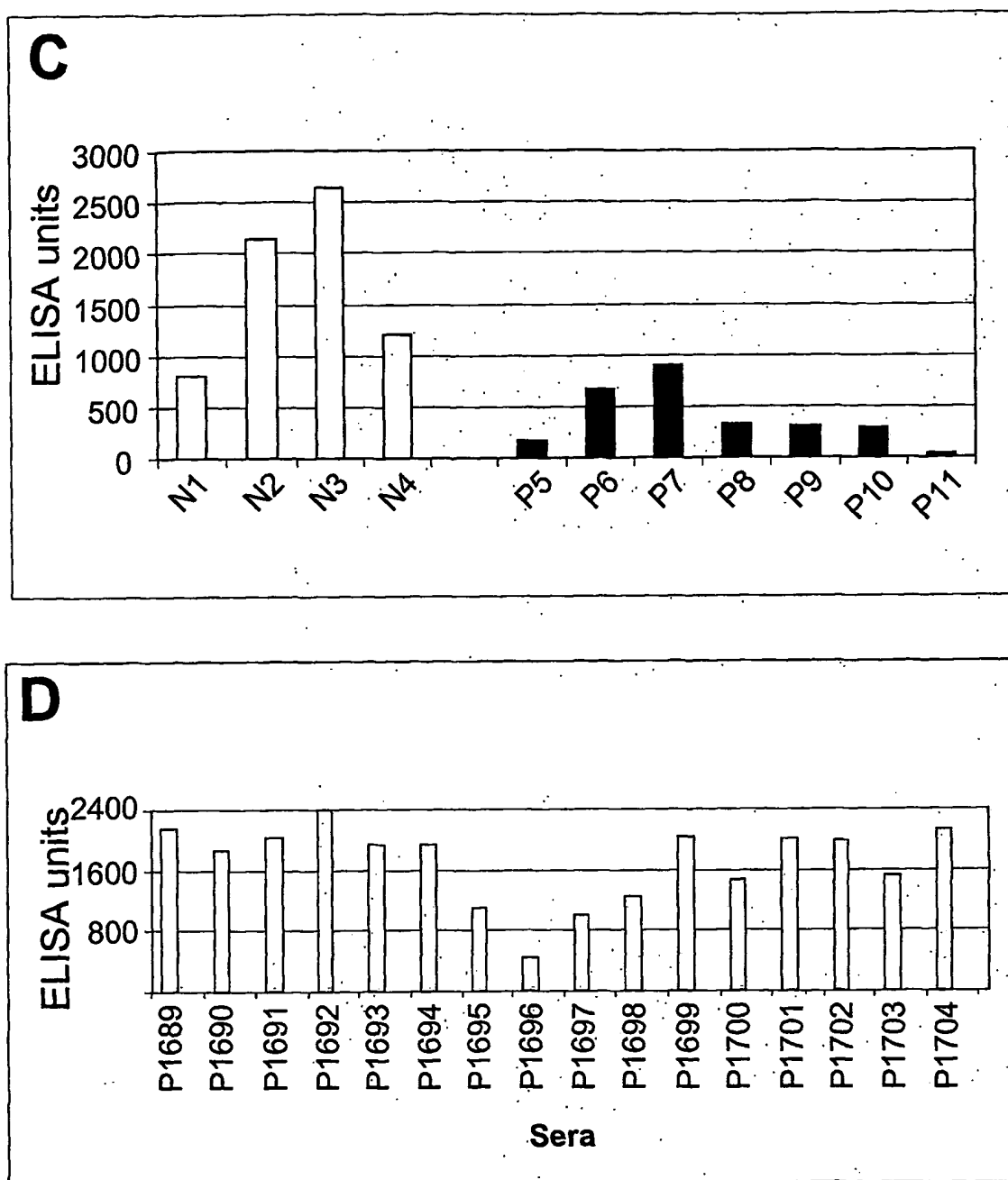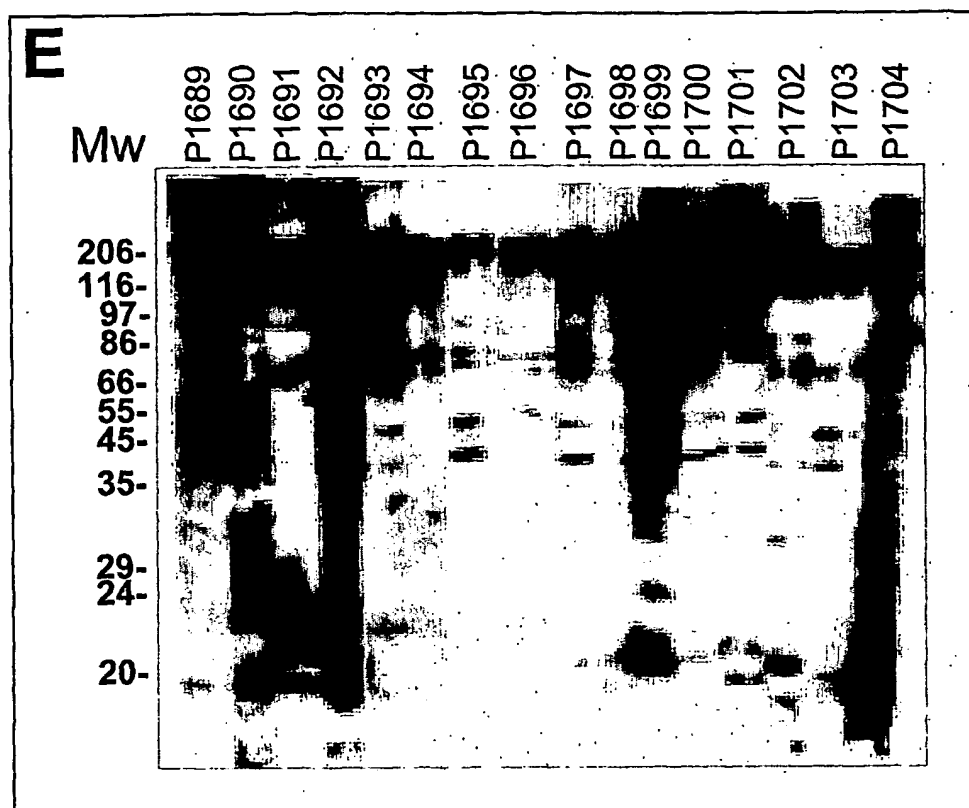
Fig. 1

Fig. 1
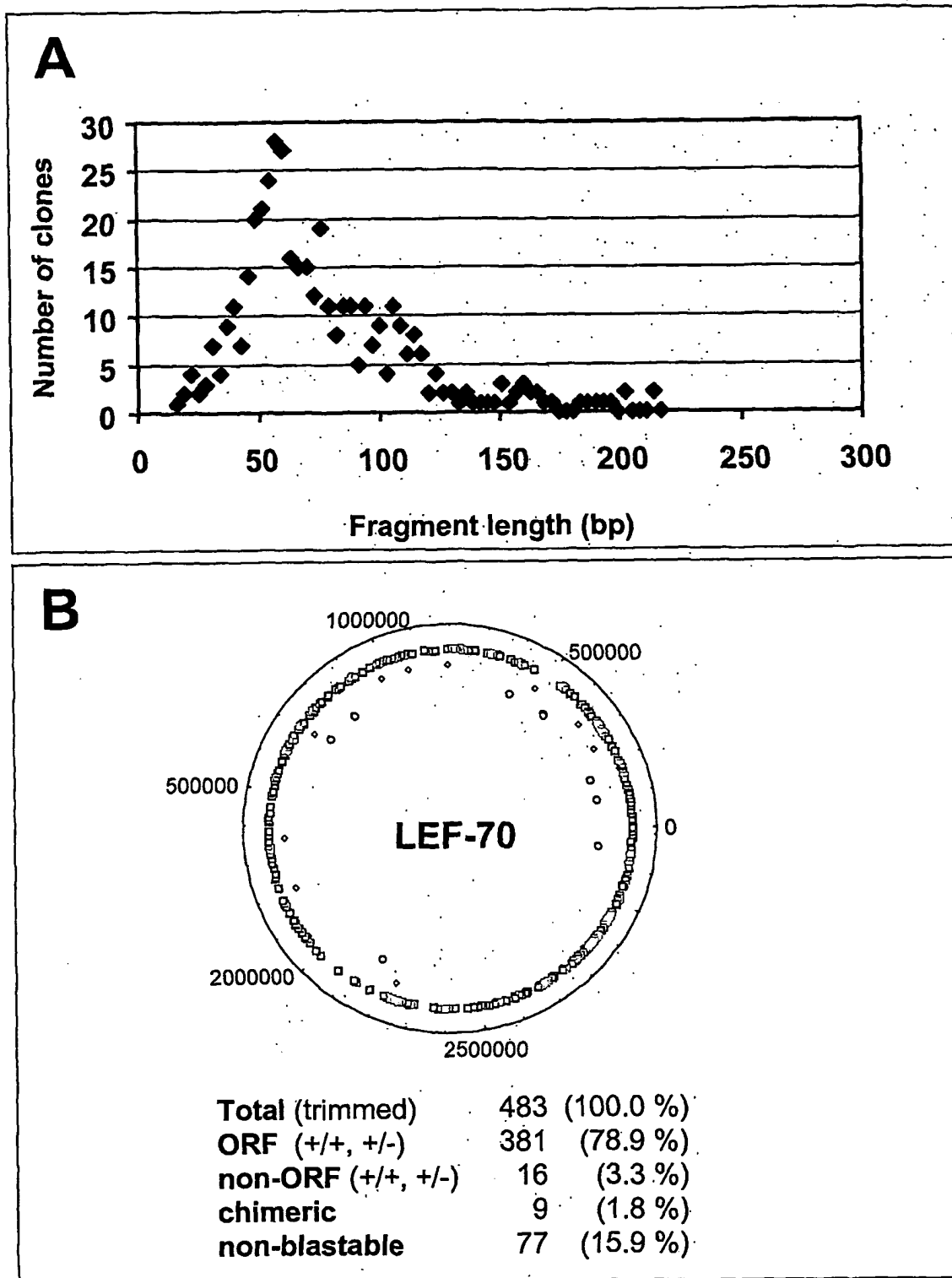
Fig. 2
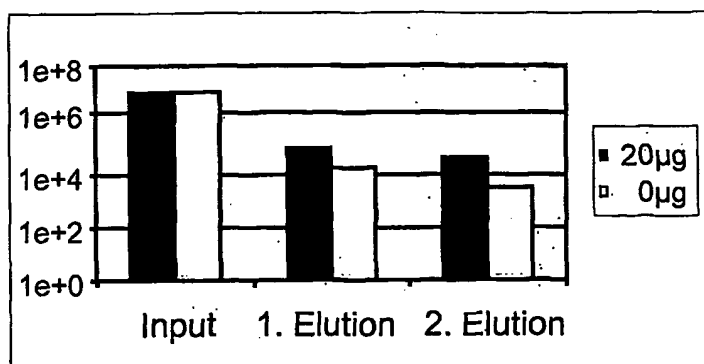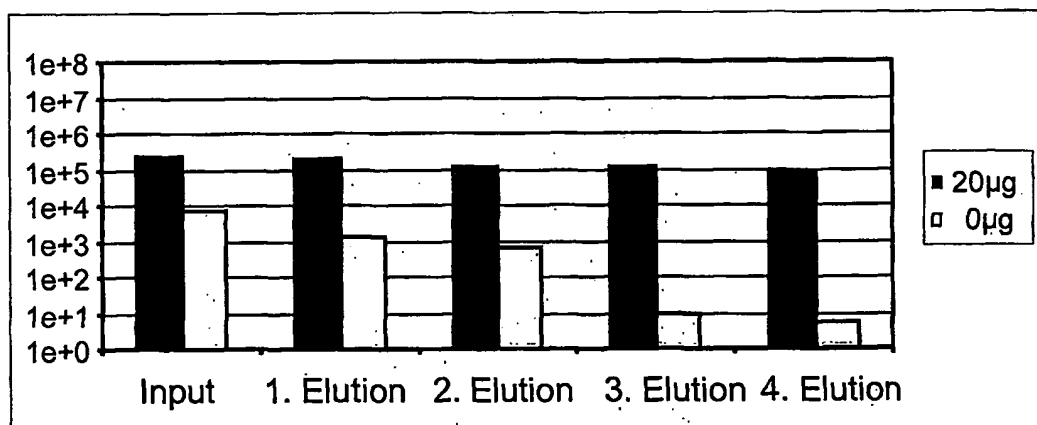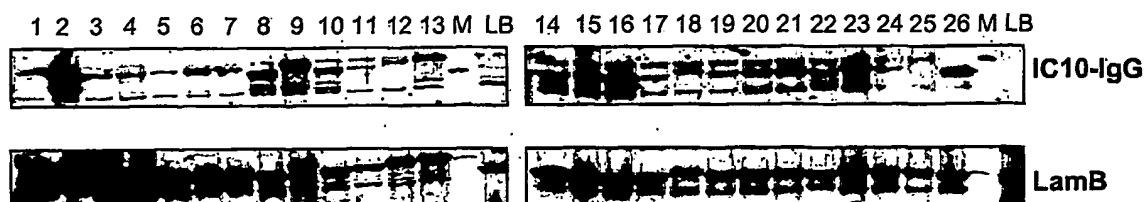
Fig. 3

SEQUENCE LISTING

```
Seq ID 1
atggtaggaa ttatcctagc aagtcacggt gaatttgctg aagggatctt gcaatctggt        60
gcaatgatct ttggcgaaca agaaaatgta aaagcagtta cgttaatgcc tagcgaaggt       120
cccgatgatg taaaagcaaa aatgcaagaa gccattgcat cgtttgacaa tcaagacgaa       180
gtactatttt tagttgacct atggggagga acaccattta accaagcgaa ctcattgtta       240
gaagatcata aagacaaatg ggcaatcgtt gcaggtatga acttaccaat ggttattgaa       300
gcatatgctt ctcgtttctc tatggaatca gcacaagaaa ttgcgacaca cattcttgaa       360
acagcaaaag atggtgtcaa agttaaacca gaagagttac aaccagcaga agcaccaaaa       420
gctgctgcga cagaagatgc acagccaaaa ggttcgttac caccaggtac agttgttggc       480
gatggcaaaa ttaaatttgt gctagcacgt attgattcac gtttattaca tggtcaagta       540
gcaacagctt ggacaaaagc aacacaacca aacagaatta ttgttgtttc tgatgccgtt       600
gcaaaagatg acttacgtaa gaaattaatc gaacaagcag ctcctccagg agtaaaagcc       660
aatgttattc caatcagtaa aatgattgaa gtagcgaaag atccacgttt tggcaataca       720
aaagcattat tattatttga aaatcctgaa gacgtcttaa aagtggttga aggcggcgta       780
gaaatcccag aagtaaacgt tggttcaatg gctcactctg ttggaaaagt cgttgtaagc       840
aaagtattat caatgggtca agaagacgtt gatacgtttg atgaattaaa agccaaaggc       900
atcaaatttg atgtgcgtaa agtgccaaat gattcaaaag cgaatatgga cgaaattttg       960
aaaaaagcaa aaaatgagtt agcaaacgcg                                        990

Seq ID 2
atgaaaaata aaagtaggca gttttttaaaa aacaactggc cttatatgtt agccagtttc        60
ttcattccat tttttaattat ggcaattatc tatttgagta ttgggatta tccaggaagt       120
agtcgcagtg ttttagccag cgatgccttt tctcaatttt caaattttca cgcaagcttt       180
aacaatgttc ttcatggaaa acaaagtctt ttctatacgt ggaatgcttc tttgggacta       240
aactacttgt ccttaatttc ctattatctt ggtggcttat ttacgcgct cgtcttttc        300
tttaacaatc aaaacatgcc tgatgcgctc tatttcctaa cgttgttgaa aattgggagt       360
gctggcttga gttttttggtt cttggcaaaa caaacctta agattcctaa atggtcccat       420
gtgactttaa gtgtgtctta tgcgctaatg tccttcattg tggcccattc agaattaatt       480
atgtggctgg atgcctttat ttatctgcca ctcgtaattt taggaattca tcgtttaatg       540
gaccaacgaa aaccgacatt attgtttgtc agttatttct tgctttcat tacaaattat       600
tattttggct ttatgattgg gttgtttcc ttcctgtact attttgcacg aacgtttaca       660
gactggcaac gatataaatc tcgaatcgtt gcgtatttta ctacttctct tttagctggt       720
ggagcatcca tgattatggt cttacctgct gtcttagatt tacgaaccaa tggtgaaact       780
ctttctgaaa ttacaacctt taaaacagaa gcgactgcct tttagatat tattatgaaa       840
aatatgattg gcgtttacga tacaaccaaa tatggctttat ttacattggg                900
ttgttgcctt tgattttttt tctcttctat ttcgttacga aagaagttcc attgaaaaac       960
aaactacttt ttggcagttt gtttgtccta ttgattgcta gctttacat cacgcccta       1020
aatttattct ggcacggtat gcatgcgccg aacatgtttc ttttccgcta tagtttcctc      1080
ttttcattct tagttatttt gcttgcaggt tatggttggg agaaatttga aaaggatgat      1140
ctcggtgtct taagcggctt gattttgatt ctttttagcaa ttttttgctct tgcaatgggc      1200
acaaaaggtg ctactagtta tacgtatgtg acttttaacct ctttcgtatt aacggcaact      1260
ttttactat tatactttt cgggattgct tttatcaat tgaaaaaagc gcctatgcaa         1320
tacttagtga ttttgttatt gctgctcgtt tctggtgaag caatcattaa cacacgggcg      1380
atggtcactg tattttaga tgattggaac tacgcatcac gcagcttgta ttctgagcct      1440
tatcctgact taaaaaatct agtcgatcaa acaaaaaaag aaaatgatac cttttaccgc      1500
ttggaaaatt taaatggtgt ttcggccaat gacggcatta actatggcta cagtggcatt      1560
agcatgtttt cttccgtacg gaaccgacat tcttcacatt acctaaacgc cttaggtttc      1620
cgctcacgag gcacgaattt aaatattcgt taccaaaata taccttgtt aatggatgct      1680
ttgatgggca ttaaatacaa tattgctgaa aacaatccaa tgaaattcgg ctttgagcga      1740
caagctgctg ctggtaaata tcagctttat cgtaacgaaa atgctttgcc gttaggattt      1800
ttagcagaca aagagattta caatgtccgg caacctctca atgataattt aggtagtcaa      1860
acaaatttac taaaatgcttt agccaatact aatgaacggt attttaccctt ttatcaacca      1920
acgatgacgc ttcaaaataa cgttgacaatt acgcaaaata cagctggcgt aactttacg      1980
gagaaacaac ataacgtagc caaagagatt tcttatacag tcaatgtgcc cgccaatacg      2040
caagcttatt tgagtttatt cccaacggat tttgctcaat tggaaagttc cactgcgacg      2100
gtcaccgtga acggctcaag tcagcaatca caaattggca ttactggcca atattataac      2160
ttaggttact atcccaaaga tacaaccgtt aacttcaaag taagttttta tggaaccaaa      2220
gcagtaagtt ttgttcagcc gcaagtggtt ggcttaaata ccaatgcatt tgaaaaagcc      2280
atttccgccg tgcaagaaaa aggcgttgat ttaacaactg gcaaacgctc tgcttcaggt      2340
acctttaccg ccgacaaaga tcaagtactc gtcactacaa ttcctttatga caaggttggg      2400
cgtgtcaaaa ttgatggcaa aaaagttacc ccaaaagcct ttaaagatgc ctttctaagc      2460
gttcctgtca gcgctggaac ccatactatt cagttctctt atttaccaga aggattgatt      2520
cctggcatcg tgttattcgt tctatgtact ggtggtttcg ttgcctacgt aaccttgatt      2580
cctgctagac gaaaccgtaa aaaagaagac aaa                                   2613

Seq ID 3
gtgaaaaagt tgagcttttaa aaaagtgaag tggggcatgc attttttaat ggctgttgcg        60
ttgatagcgc caagtgttac tagtacggca tatgcagtag aaacaacgag tcaacaaagt       120
tcagaagcag taacaagtac caccgattca agtagaaaac aagaaccagt cattacacag       180
```

```
gaaacaacag acatcaaaca agaagcacca aatcaggcta cgagtgacag tgtcaagcag    240
tcacaagaaa ccacagcacc aacagagacg acgaatttag aaacgtcaat cgctgaaaaa    300
gaagaaacga gcacgccgca aaaaataaca attttaggta cgtcagatgt tcatggtcaa    360
ttatggaatt ggtcttatga agatgataaa gaactaccag ttggtttgtc ccaagtaagt    420
acagtcgtta accaagtccg ggcacaaaac ccagcaggca ccgtttttaat tgataatggc    480
gacaatattc aaggcactat tttaacagat gacttgtata ataaagcgcc tttagtgaat    540
gaaaagaccc atccaatgat caccgccatg aatgtgatga agtatgatgc aatggttttg    600
ggaaatcatg agtttaattt tggtttaccg ttaatcaaaa aaattcaaca agaagccact    660
tttccaatct tgtctgcgaa tacctacaat aaggaagatg gtcttcgttt tgttgaaggg    720
actaccacga aggaacttga tttttaatcaa gatgggcagc cagatttaaa agttgggatt    780
atcggcttaa caattccgca cattcctttg tgggatggcc ctcgtgttac ttcgcttaat    840
tttttacctt tgaaagaaga agcagaaaaa gcagttactg agttgaaagc taacgatcag    900
gctgacatta ttgttgcctc gattcatgcg ggacaacaaa atagtgatcc ggctgccagt    960
gccgaccaag taattgaaaa tgtcgcgggg attgatgcgt atattctggg tcatgaccac   1020
ctttcttta ccaagcaagg agcagcgccg aatggaaaaa ctgtaccggt aggggggaccg   1080
aaagatacgg ggacagaagt tgtcaaaatt gatctttcag ttgctaaaaa tgccgataag   1140
tgggaagtgc aagaaggtac agcaacgatt gtaccaacaa cgaatgttcc agcagatgaa   1200
gcagttaagg cagcgacaaa agaataccat gaaaaaacgc gagcgtttat tcaggaggag   1260
atcggcacag caacagctga ttttttacca aaacaagaaa ttaaaggaat tcccgaagca   1320
caattacaac caacagcgat gatttcttta attaataacg ttcaaaaaga agtaacgggc   1380
gcacaattaa gtgcggcagc gctgtttaaa tacgcagta aattacctgc ggggaagatt   1440
tcctatgcca cgattttga tatctacaaa tacccgaata ccttagtgag tgttcccatt   1500
aacggtgaaa acttactgaa gtatttagaa aaacaagggg cgtactataa ccaaacacag   1560
ccagatgatt tgaccattag ttttaatcca aacattcgtg tatataacta tgacatgatt   1620
tctggagtgg actacaagat tgacatttca aaaccagtgg gtgaacgaat tgtagatgcg   1680
aaaattgacg gccaaccgct ggatcctgcc aaagaatata cgattgctat gaataattat   1740
cgttacggcg gtttagctag ccaagggatt caagtagggg aacctattaa aaattctgat   1800
ccagaaacct tacgaggaat gattgttgat tatattaaga aaaaggaac tcttgatcca   1860
gaacaagaaa tcgaacgaaa ttggtcaatt attgggacaa attttgatga aaaatggcgt   1920
gccaaagcaa tcgaattagt gaatgacggc actcttcaaa ttccgacttc tcctgatgga   1980
cgtacaccaa acgccgcgc tattacgaaa caagatgtcc gtaatgcggg ctttgattta   2040
gataatgcat ataccattat gcacacaaat gacgttcatg gccgactaga agcagggaaa   2100
ggcgaattag gtatggcgcg tctaaaaacc tttaaagacc aagaaaaccc aaccttgatg   2160
gtggatgcag gggatgtttt ccaaggatta ccaatctcca atttctccaa aggcgcggat   2220
atggccaaag caatgaatga agttggttat gatgccatgg cggtgggaaa tcacgagttt   2280
gattttggtt tagagattgc actaggttat aaagaccaac tgaattttcc gattttatct   2340
agtaatacgt attacaaaga tggcagtgaa cgggttttg atccgtatac aatcgtagaa   2400
aaatccggga aaaagtttgc cattgtaggt gtgacgaccc cagaaacagc aacgaaaaca   2460
cacccgaaaa acgtagagaa ggtgacattt aaagacccga ttccagaagt agaagcagtg   2520
attaaggaaa ttaaagagaa gtacgcggat attcaagctt tcgtggttac tgggcattta   2580
ggcgtagatg aaacgacgcc gcatatctgg cgtggtgata cgctagcaga aaccccttagt   2640
caaacatatc ctgagttaga tatcactgtg attgatggac attcgcatac agccgtcgaa   2700
agtggcaaac gttatgtcaa agtgatctat gctcaaacag gtaattatt aaataatgtt   2760
gggatcgtca cagcaccaga gagtgaacca actaagaaaa caacaaaatt gatttcagca   2820
gcagagctgc tagaattgcc agaaaaccg gcagttaaag ccatcgttga tgaagcacgt   2880
acgaatttta acgctgaaaa tgaaaaagta attgtcgatt atattccatt cacattggat   2940
ggacaacgag aaaatgtgcg cacacgagag accaacttag ggaatttgat tggtgatgcg   3000
attatgtcat atggccaaga cgcgtttagc caacctgctg attttgcagt aactaatggt   3060
ggcggcattc gcgctgatat taaacaaggt ccaattaaag ttggggatgt cattgctgtg   3120
ttacctttg gcaatagcat tgcgcaaatt caagtaaccg gcgccaagt taaagaaatg   3180
tttgaaatgt ctgttcgttc gattccacaa aaagatgaga atggcacaat tttactagat   3240
gatgctggcc aaccaaaact tggcgcaaat ggtggtttcc tacatgtttc aagctccatt   3300
cgtatccact atgattccac aaaaaccaggt actcgcttgg ctagtgacga aggcaatgaa   3360
acaggacaaa cgattgtcgg tagtcgcgta ttaggaatag aaattaaaaa tcggcaaaca   3420
caaaagtttg aaccattgga tgagaagaaa caataccgga tggctaccaa tgatttctca   3480
gctgctggtg gtgatggtta cgatatgcta ggtggtgaac gagaagaagg gatttcacta   3540
gattctgtct taattgaata cttgaaaagt gcaaccagct tgcggttgta tcgtgcagca   3600
acgacgattg atttagcaca atataaagaa ccattcccag gcgaacgaat tgtttctatt   3660
tcggaagaag cttacaaaga gttaatcggt ggaggagaga cgccaaaacc agatccaaaa   3720
ccagacccga aaccaacacc agaaacacca gtagcaacca ataaacaaaa ccaagcggga   3780
gcaagacaga gcaatccatc cgtaacagag aagaaaaagt atggcggctt tttacctaaa   3840
acgggtacag aaacagaaac gcttgcatta tatggtttac tgttcgttgg actttcttct   3900
tctggctggt atatttataaa acgacgtaac aaagctagt                           3939
```

Seq ID 4
```
atgaatcaac agactgaagt aaaaaaacgt tttaaaatgt ataaggcaaa gaaacactgg    60
attgtagccc ctattctttt tttaggtgta ttaggagctg taggattagc tacagataat   120
gtacaagccg cggaattaga tacgcaacca ggacaaacga cggtgcaacc cgataatccc   180
aacccgcagt cacgaaatga aacacttaaa acggcagtat ctgaagaagc agcattacaa   240
aaagacacta cttctcaacc aaccacagca gaagaagtag tgccgaaagg aattgctgct   300
gaacaaagtt cagctacctc aaatgatacc acaaacgtcc aacaaccaac agcagaagca   360
```

```
gaaaaatcag cacaagaaca accagtagtc agccctgaaa caaccattga acctctaggg     420
cagccaacag aagttgcacc agcagaaaat gacgctaata aatcaacgtc aattcctaaa     480
gaatttgaaa caccagacgt tgacaaagca gttgatgaag cgaaaaaaga tccaaacatt     540
accgtcgtgg aaaaaccaac tgaagactta ggaaatgttt cttctaaaga tttagctgca     600
aaagaaaaag aagtagacca actacaaaaa gaacaagccc aaaagattgc ccaacaagca     660
gctgaattaa aagccaaaaa tgaaaaaatt gccaaagaaa atgcagaaat tgcggcaaaa     720
aataaagcgg aaaaagagcg ctacgaaaaa gaagtggcgg aatacaacaa acacaaaaat     780
gacaaaggct atgtgaatga agcaatcagc aaagacttag ttttttgattc aagcattgtg     840
actaaggata ctaaaataga caagattaca ggtggcaagt ttattaaagc ttccgatttt     900
aacaaagtaa accaaggaca atccaaagat attttttacaa aattaagtaa ggatatgaat     960
ggaaaagcaa caggcaactt ccaaagctca aaagtggctg ctgttgaatt tggccctaaa    1020
ggaggatacg cggttctttt agaaaaaaac aaaccggtta atgtcacgta tacaggatta    1080
aatgctagtt atttagaccg aaaaatcaca aaagcgagt tcatttatga acttcaatct    1140
gcaccaagtc aaagcggcac cttaaatgca gtattttcga atgatcctat tattacagca    1200
tttgttggta ctaaaaatgc caatgggaag gacgtaaaag ttcgcttaac cattaaatta    1260
tatgatgcta atggcaaaga agtttttacca gataaagacc atgcctttgc ctatgcgtta    1320
tcgtcattaa actctagtct aggaacaaat tatagtgtag aacatgcgga atttgtttca    1380
gactttggct caaaaaatga gttcaagtat attaatggtt cctacgtaaa gaaacaagcg    1440
gatgggaaat tctactcaac cgaagacctt gattatggga caggtccttc aggactaaaa    1500
aatagtgatt gggacgctgt aggtcataaa aatgcgtact atggttctgg tgtaggtcta    1560
gctagagagg gtggacgtat ttctttctct tttggtatga caaccaaagg aaaagtcaac    1620
ctatctggtg cgcaatggtt tgcctttagt accaatttaa atgcgaaatc aattagacca    1680
taccaaaaga aagggaatcc aaaagaacca gaaaaagcaa caattgaatt caatcgatac    1740
aaagccaatg tcgttcctgt tcttgtgccg aataaagaag tcactgatgg tcagaaaaat    1800
atcaatgatt taaatgtgaa acgaggcgat tctttacaat acattgtgac aggggatacg    1860
acagaacttg ccaaagtaga tccgaaaaca gtgacaaaac aagggattcg ggatacctt     1920
gatgcagaaa aagtgacgat tgatttatcc aaagtgaagg tttatcaagc agatgcaagt    1980
ctaaacgaga aagacttaaa agctgttgct gcagcgatta attcaggaaa agctatagac    2040
gtgactgctc cttatgttct taatttagat caaaacaccg tcacagcaat gatgaaaacc    2100
aacgcagacg gttccgttgt tttagcaatg gggtataaat atttacttgt cttgccgttt    2160
gtagtgaaaa atgtagaagg cgattttgaa aatacagctg ttcagctgca aaatgatgta    2220
gaaacagtaa caaatacagt gattaaccat gtgccggtta gtaatccttc caaagatgta    2280
aaagcagata aaaacggtac agttggcagt gtttctctac atgataaaga tattccgtta    2340
caaacaaaaa tttattatga agtgaaatct tccgaacgtc cagctaacta tggcggaatt    2400
accgaagaat ggggcatgaa tgatgtcttg gacacgaccc atgatcgttt cacaggtaaa    2460
tggcacgcta ttacaaacta tgaccttaaa gtaggggaca aaacgttaaa agcaggaaca    2520
gatatttctg cctacattct tttagaaaac aaagacaata aagacttgac gtttacaatg    2580
aatcaagcat tattagcagc gttaaatgaa ggaagcaata aagtaggcaa acaagcttgg    2640
tctgtatatc tggaagtcga acggatcaaa acaggtacg tagaaaatac gcaaacagaa    2700
aactacaaca aagaacttgt tcgttctaat acggtggtga cgcatacacc tgatgatcca    2760
aaaccaacca aagccgttca caacaagaaa ggggaagata ttaaccatgg aaaagttgct    2820
cgtggtgatg ttctttctta tgaaatgacg tgggacttaa aagggtacga taaagacttt    2880
gcttttgaca cagtcgatct tgcgacagga gtttctttct tcgatgatta cgacgaaacg    2940
aaagtgacac caatccaaaga cttacttcgt gtcaaagatt ctaaagggc agacattacg    3000
aaccagttca cgatctcttg ggacgatgcc aaaggcacgg tgacaatctc tgccaaagac    3060
ccacaagcct ttattctagc gtatggtggg caagaattgc gtgtaacgct ccctacaaaa    3120
gtcaaagcca atgtttctgg tgatgtttat aattcagcgg aacaaaatac atttggtcaa    3180
cgaattaaaa caaataccgt tgtcaaccat attccaaaag tgaaccctaa aaaagacgtg    3240
gttattaaag ttggtgacaa acaaagtcaa aatggcgcca caatcaaatt aggggagaaa    3300
ttcttctatg aatttacaag tagtgacatt cctgcagaat acgctggtgt tgtgggaagaa    3360
tggtcgatta gcgataaact agacgtaaaa catgacaaat ttagtggcca atggtctgtg    3420
tttgccaatt ctaattttgt tttagcagac ggaaccaaag tgaataaagg gacgacatt     3480
tcgaaactat tcacgatgac ctttgaacaa ggggtagtta aaatcacggc cagtcaagcg    3540
tttttggatg cgatgaatct aaaagaaaac aaaacacgttg cgcactcatg gaaagcgttc    3600
attggtgtag aacgaattgc ggcaggagac gtttacaaca caatcgaaga atctttcaac    3660
aatgagaaga ttaaaactaa tacggtagtg acacatacgc cagaaaaacc acaaacacca    3720
ccagaaaaaa cagtcgattgt accaccaaca ccaaaacaac cacaagcacc agtagagcca    3780
ttagtggtag aaaaagcaag tgtggcaccc gaactacctc atacaggtga aaaagaaaac    3840
accctattat ctgtactagg tgccggaatg ctagttggtc tggcttggtt tggtttgaaa    3900
aaacgtgaag tgaaa                                                    3915


Seq ID 5
atgaaacgaa ttggttatgc acgcacaact attattgaag acgatttaaa aaatcaacta      60
accactctcc aatcgtttgg ttgtgatgat attttccaag agacatttga tcctcaagcg     120
gagatcagtg tcttagacga agtcgaaaaa cttctttctg ctggggatac acttattgtt     180
tgtaaattgc accatttagg taaaacaacc cgccaattaa cagactttat gaaaatgtta     240
aaggaaaaac aagttgattt tgtcagcatt tcagaaggaa ttgatactca ccttccaaca     300
ggtgaagcct atttccaatt aatggagagc ttatctgcga tggaatgtgc actaattaaa     360
gaacgaacac tcgtcgggct ccacaaagct cgtgaaaacg cgtgaaagtagg tggacgtcca     420
aaaatcgatg gacgcaccgt ccgcaaaatc cgtgcattat attatgaaaa caaagaaaca     480
atccaattta tttctaataa atgcggcgtt tcggtgggca cttgttataa gtacatcaat     540
```

```
ttacctgaga cagatgtcga gcggctgtat tcc                                      573
```

Seq ID 6
```
atgagtaaaa aagaaataaa tcaagtagtt gccagtagct atcaattgta tattaatgga        60
gagtggacaa caggtagtgg taacaaaatg attgctagtt acaatcctag taatgtcgaa        120
aaattagcag aatttgtaga tgccacaaat gcagatgtgg atcgagctgt agaagcagcc        180
caagaagcgt ttcagacatg gaaagatgtt gacgttgtaa caagaagcaa tctttttgttg       240
aaaattgctg atttgattga agaaaatcaa gaacatttgg ctatggtgga gactttagat        300
aatggaaaac cgcttcggga aacgcaatcg attgatgtcc ctgccagtgc agaccatttt        360
cggtattttg ctagtgtgat tcgtgagaa gaaggatctg tcaaagaatt tgataaggat         420
acgttatcca ttgttgtgaa agaacccatc ggtgttgtgg gtcaaattat tccgtggaat        480
ttccccttat taatgggtgc ttggaaatta gcaccagcgt tggcagcagg taacacggtt        540
gtgattcatc cgtcctcgag cacatcatta agtctgttgg aattgtttaa aattttgat          600
caagtcttgc cgaaaggagt agtgaattta atcaccggtc gtggttctga ttcaggaaat        660
tatatgttgg cacatccagg gtttgataag ctagctttta caggctcaac agaggtgggg        720
tacactgtcg ctaaagcggc ggccgaccgc ctaattccag ccactttaga acttggcgga        780
aaatcagcca acattatttt tgaagatgcc aattgggaac gtgcattaga aggcgtgcag        840
ttagggattt tattcaatca agggcaagtt tgttgtgctg ggtctcgtgt gttcgttcag        900
tcaggtattt atgatcaatt tgtagaagct ttaaaggaaa agtttgaaca agtgaacgtt        960
ggtttcccgt gggaaaaaga tgttgaaatg ggcgctcaga tcaatgagca tcaattggaa        1020
gaaatttttaa aatatgtcga aattggtgtg aaggaaggag ctacgctgat tactggtggg       1080
caacgtttaa cagaaaatgg gctagacaag ggggcgtttt tagcacctac gttattagcg       1140
aatggtacga atgcaatgtg tgtggcccaa gaagaaatct ttggtcctgt tgcaacagtg        1200
attaaatttg aaacggaaga agaagtcatt cgtttagcta atgattctga atatggtcta        1260
ggtggtgccg tcttttctca agatatcaat gtggcattac gggttgctcg tggtgtacgg        1320
acaggtcgaa tgtgggtcaa cacatacaat caattgcctg cgggcgcgcc atttggcggt        1380
tataaaaaat caggaattgg tcgagaaacg cataaatcaa tgctagatgc ttatacgcaa        1440
atgaaaaata tttacattgt gacaaaagaa gaagcagatg gactgtat                    1488
```

Seq ID 7
```
atgggaaaat atcatgaact agcagaaaaa attgtaaaaa atgtcggtgg gcaagagaat        60
attaatagtt tgacacattg tatcacacgt cttcgcttca aattaaaaga tgagagtcaa        120
gcgaatgatg atgtcttaaa aaatatggac ggtgtggtaa ctgttatgaa aagcggtgga        180
caataccaag tcgttattgg gaaccatgtt ccagccgttt acgaagaagt agttagtatt        240
gcaggcttgt caggagaacg tgaagaagaa gcttccagtg ggaatttatt tgatcgctta        300
atagatattt taagtggttg tttccaaccc tttttagggg cgttagcagc ggctgggatg        360
gtcaaagggt taaatgcctt attggttttc ttaaagtctct atacggctac gtcaatggact       420
tataccatgt taaatgggat tggcgatgca attttctact ttatgccagt tattttaggt        480
tatacagcag ctaaaaaatt ccggttgcat ccaatggtag ggatcgtgat tggcgcagct        540
ttatgttacc caacgattca aggaagtgca ttacaaacgg cctttgaaac aacagcgggt        600
gccggagccg cagcaccgta taatttattc ggtttacctg cctataatac atttatgggg        660
attccttggg tcggtgcgaa ctatacaagt agtgtcgtac caatcatctt cattattgct        720
tttgcggcac aagtgtttaaa cgtattattc cagaagtttg tcaaacgttc                   780
ttagtaccgt tttttcgtctt gttgattgcc ttaccaattg gtttcttagt aattgggcca       840
atcgttagca tgctaacaga tttattaagc gctggcttta cagcattaat gagtttctca        900
ccagctttgt atggtttgat tcttggtttc ttctggcaag tcttagttat ttttggttta        960
cattggagcg ttgtgccatt agctattatg caagtgacac aagaaggttc aagccaagta        1020
ttaacgggat catttgcagc tagttttgcg caaacagcgg ttgtttagc gatgttcttc         1080
aaattgaaag ataaaaagtt aaaagctttg tgtccgccag caattatctc tggtattttt        1140
ggtgtaacag agccagcaat ttatggaatc actttaccaa aaaaatggcc atttatttat        1200
tcaatgattg gtggcgctgt cggtggttta tatttgatga tcaacaatgt gacagcttat        1260
acaatgggtg gtttaggaat ttttggtgtg ctaaacttta tcaatggtga tgatgcaagt        1320
gggatgatcc agtcctttat tgcaattgcc attgcagcgg tcgttggttt tggtctaaca        1380
ttcttcttct ggaaagacaa tacagttgag gaagaagaag tgattattga taaaacaacg        1440
attaaaaaag aaaatattac aagtccagtc aaaggacgag tgttatcttt aaaaaatgca        1500
gaggatcctg cttttgtcaa tggagcatta ggaaatgggg tcgtgattga accaacggaa        1560
ggtaaagtgg ttgcacccct tgatggaacg attgtcacac tattcccaac aaaaacatgca       1620
ttaggtttga tttcagataa cggcacagaa ttattgattc atattgggat tgatacagtt        1680
caattagaag gcgaaggctt tgaagctttt gtaaaacaag gtgaccgcgt gaaaaaaggt        1740
caaacattag taacctttga tttagaagga attaaaaaag ctgggtttag tacacaaatt        1800
ccaatcgttg tgacgaatac agcggattat ttagatattt tggaagtcgg cagtaatgaa        1860
gtttcaacaa gtgatgattt gctgaccgct ttaata                                  1896
```

Seq ID 8
```
atggaaagta aaacgtttga tattgaaggg atgagctgtg cttcctgtgc ccaaacgatt        60
gaaaaggcta cggcgaagtt acctgggat gcaaaggcaa gtgtcaattt agctacgaa          120
aaattaagtg tcacgtatga tcagacggaa gtcacagaag aagaaattaa ggaagctgtg        180
tcagatgctg gctacaaggc gattagccca gcccagcaac gaacgtttgc gattgaagga        240
atgagttgtg cttcctgtgc ccaaacgatt gaaaaagctg tgaaccaatt atctggtgtt        300
caacaagcga ttgttaattt agccacggaa aaactggtgg ttagttatga cgaccatcaa        360
gtaacgagtg ctgaaattat caaagcagtt acggatgcag gttatcaagc gacagaagaa        420
```

```
gtcgctgcag gtgcaacagc ggaccaagat cgtgaaaaga aacaaaaaca tattgctgag   480
atgtggcaac gtttttggat ctcagcggtg tttacagtgc cccttcttta tattgcgatg   540
ggacatatgg taggcttacc attgccggat ttttttgaatc caatgacaca tgcgatgacc   600
tttgcaatgg ttcaattgat tttgacttta cccgttctat atgtcggcag agaattttc   660
acagttggtt ttaaagcatt gtttaaaggg catccgaata tgtttttctt ggtggcttta   720
ggaaccagcg ctgctttgt gtatagttta tatggcacag tcatgatatt tttaggagac   780
acaagtttta caatggctct ttattatgaa tcagcgggtg ttattttaac cttaattaca   840
ttaggaaaat atttcgaagc tgtttcaaaa gggaaaacct cagatgcgat aaaaaaatta   900
atgggtttag cgccgaaaac tgctcatatt ttacgtgatg gcgtagaaat tgaagtgcca   960
gttgatgctg ttcagttgga tgacattgtg attgttcgcc caggagataa aatccctgtt   1020
gatggagtga ttgtttcagg cagtagttct gtggataag ccatgttaac gggcgaaagt    1080
ttgcctgttg aaaagaaagt cggtgatgca gtaattggtg ctagcatcaa taaaaatggc   1140
agctttcaat tcaaggcgac aaaagtggga aaagagacgg cccttgccca aatcattcaa   1200
ttagtggaag atgcgcaagc tcaaaaagcg ccgattgcac aattagctga taaaatttcg   1260
ggtgtttttg tcccaatcgt cattggtttg gctgttcttt ctggcttggc ttggttcttt   1320
ttaggacaag agtcatggat tttcgctttg acaattacta tttctgtgtt agtgattgct   1380
tgtccttgtg cgctcggctt ggcaacgccg actgcgatta tggtggggac aggtaagggc   1440
gctgaaaatg gcgttttaat taaaagtggg gatgcttatc aaaccactca caaattcaa   1500
acaatcgtgt ttgataaaac agggactatt acagaaggca aaccggtggt tacagatatt   1560
ttagttgctg attcagcact ttcagaagct gagttactaa cattagcagc ttctgcagaa   1620
caaggatcag aacatcctct aggtgaagca attgtgggag cggctaaaga acgtcaattg   1680
ccattagcag aaggaagcga ttttttcagcg attccaggtc acggcatccg tgtgacagtc   1740
aacgaacgag tcctactttt gggcaatata aaattaatga aagaagaagc aatcgagcta   1800
tcaacgtttg ttcaacaagc ggatcgtttg gcagaacaga ggaaaacccc aatgtttgaa   1860
gcaaaagatg gttcgtttgc gggtatcatt goggtagcag atactgtgaa ggatagcagt   1920
caaacagcca ttgcacgttt gcataaaatg gggatcgaag ccgtaatgat tacgggagac   1980
aacaagcgta ctgcagaggc gattgcgaaa caagtaggta tcgatcgtgt cttaagtgaa   2040
gtttttaccag aagataaagc gttagaagtg aaaaaactac aagcagaagg caaaaaagtt   2100
gcgatggtgg gtgacgggat taatgatgca ccagcacttg cccaagcaga tgtggggatt   2160
gctattggtt caggcacaga tgttgcaatg gaatcggccg acattgtgct aatgcgtagt   2220
gatttaatgg atgtgcctac agcagttgaa ttaagtaaag cgaccatcaa aaacatcaaa   2280
gaaaatcttt tctgggcatt tgcctataat acattaggta ttccagtggc tatgggtgtg   2340
ttacacttat ttggtggacc tttgttaagt ccgatgattg cagcagcagc tatgagtttt   2400
agttcagtct ctgtcttact aaaacgcctta cgattgaaag gatttaaacc ttcgacggtg   2460
aaaagaacta gtggaagtca aaaa                                         2484
```

Seq ID 9
```
atgaaaaaga aaatttttagc aggagcgctt gtcgctctgt tttttatgcc tacagctatg    60
tttgccgcaa aaggagacca aggtgtggat tgggcgattt atcaaggtga acaaggtcgc   120
tttggctatg cacatgataa attcgctatt gcccagattg gaggctacaa tgctagcggt   180
atttatgaac aatacacata taaaacgcaa gtggcaagtg ctattgccca aggtaaacgt   240
gcgcatacct atatttggta tgacacttgg ggaaacatgg acattgcgaa aacaacaatg   300
gattactttt tgccacgtat tcaaacgcatt aaaaattcca tcgttgcatt agattttgaa   360
catggagcgt tggctagtgt tccagatgga tatgtgaggat atgtaagttc agatgccgaa   420
aaagcagcaa atacagagac aattttgtac ggtatgcgca gaatcaaaca ggctggctat   480
actccaatgt attacagcta taagccattt acactaaatc atgtaaacta tcaacaaatc   540
atcaaagagt ttcctaactc tttatggatt gctgcgtatc ctatcgatgg tgtgtcacca   600
tatccattgt atgcttattt cccaagcatg gatggtattg gtatttggca attcacatcc   660
gcttatattg caggtggtt agatgtgaac gtagattaa caggaattac ggatagtggt   720
tatacagata ccaataaacc agaaacggat acgccagcaa cagatgcagg cgaagaaatt   780
gaaaaaatac ctaattctga tgttaaagtt ggcgatacog tcaaagtgaa atttaatgta   840
gatgcttggg caactgggga agctattccg caatgggtaa aaggaaacag ctacaaagtg   900
caagaagtaa ctggaagcag agtattgctt gaaggtatct tgtcatggat tagcaaaggt   960
gatattgaat tattgccaga cgcaacagtc gtccctgata agcaaccaga agcgactcat   1020
gtggtacaat acggagaaac attatcaagt attgcttatc aatatggaaa agacttatca   1080
acgttggcgg cattaaatgg attggctaat ccaaatctta tttatcctgg tcaagttttg   1140
aaagtcaatg gatcggcaac aagtaatgtc tacacggtta aatacggcga taatttatct   1200
agtattgcag caaaacttgg cactacttat caagctttag ctgcattaaa cggattagca   1260
aatcctaact tgatttatcc aggtcaaaca ttgaattat                        1299
```

Seq ID 10
```
ttgccgatga ttattaaaca ggaacaattc ataccaaaag atacagtaga aacaacgatt    60
gacttattaa ttagaaattt aacaacgatt aaagataaca caggtgaatt tctgttagat   120
tttgatggtt taaaagttga tgacaaaagt tggacgattt ggaactggcc gcaaggcgtt   180
gggttgtatg gtatatataa gaattatcga aatactaaaa gtgaaaaagc tttacaagta   240
gttaacgact ggtttgaagg tcggatgcaa gaaggcgcgc cgcctaagaa tgtgaatacg   300
atggctccgc tttttaacaat ggcttatcta tatgaagaca ctaaagactc taaatatata   360
ccttacttag agcaatgggc ggaatgggtg atggaagaaa tgcccccgaac gaatgaaggt   420
ggcttgcaac atgcgacata tggaccagaa aataaaaatc aattcatggga tgatacgtta   480
atgatgactg tcttaccttt ggcaaaaaat ggtaagttgt tgaatcgact cgattatttg   540
gaagaagcaa agcatcaatt tttaatccat attaaatatt tacaagataa aaaaagtggt   600
```

```
ctatggtatc acggttggac ttttgaggga aatcataact atgcggaagc actgtgggca    660
agaggaaact gctggattac gattgctatt ccagaaatta ttgaaatctt agagttgcca    720
aaaggagata gtttacggga atttttattg agtaccctca atgcgcaagt ggcagcttta    780
gcgaaatacc aagatgaatc tggtttatgg catacattaa ttaatgattc aaattcgtat    840
ttagaatctt ctgctacagc gggattcgct tatgggattt taaaagcggt tcataaaaaa    900
tatatttctt ctgaatatga agaagtggca aacaaagcaa ttgctggctt actaaatgag    960
attgatgaaa cgggagaagt acaacatgtg tcagttggta caggaatggg tgataatta   1020
gatttttatc gcacaattgg aatgacagcg atgccttatg gtcaatcatt aacaatccta   1080
tgtttgactg aattgcttgt ttcttattgc                                   1110
```

Seq ID 11
```
atgaaagcaa aaaaacaata taagacatat aaagctaaga atcactgggt aactgtccct     60
attctttttc taagtgtgtt aggagccgta ggattagcta ctgataatgt acaagccgcg    120
gaattagata cgcaaccaga aacaacgacg gttcaaccca ataaccccga cctgcagtca    180
gaaaaggaaa cacctaaaac ggcagtatct gaagaagcaa cagtacaaaa agacactact    240
tctcaaccga ccaaagtaga agaagtagcg ccagaaaata aaggtactga acaaagttca    300
gctaccccaa atgataccac aaacgcgcaa caaccaacag taggagctga aaaatcagca    360
caagaacaac cagtagtaag ccctgaaaca accaatgaac ctctaggca gccaacagaa    420
gttgcaccag ctgaaaatga agtgaataaa tcaacgtcca ttcctaaaga atttgaaaca    480
ccagacgttg ataaagcagt tgatgaagta aaaaaagatc caaacattac cgttgttgaa    540
aaaccagcag aagacttagg caacgtttct tctaaagatt tagctgcaaa agaaaaagaa    600
gtagaccaac tacaaaaaga acaagcgaaa aagattgccc aacaagcagc tgaattaaaa    660
gccaaaaatg aaaaaattgc caaagaaaat gcagaaattg cggcaaaaaa caaagcagaa    720
aaagagcgtt atgaaaaaga agtcgctgaa tacaaccgga ataagaacga aaacagctat    780
gtcaatgaag cgattagtaa aaacctagtg ttcgatcaat ctgtcgtgac gaaagacact    840
aaaatttcgt cgattaaagg cggaaaattt atcaaagcaa ctgatttaa taaagtaaat    900
gcaggggatt caaagagatat ctttacaaaa ttacggaaag atatgggcgg gaaagttact    960
ggcaacttcc agaattcctt gtaaaagag gcaaatcttg ggtctaatgg tgggtatgcg   1020
gttctttag aaaaaaataa accagtgaca gtgacctata caggactaaa cgctagttat   1080
ttaggacgta aaattacaaa agcagaattt gtttatgaac tacaatcctc accaagccaa   1140
agtggaacgt taaatgcagt attttcaaac gatccgatta tcacggcttt tattggtaca   1200
aacagagtca atggtaagga tgttaaaaca cgcttaacga ttaagttctt tgatgcgtca   1260
ggtaaagaag tactaccaga taaagatagt ccatttgcgt atgcgctgtc ttctttaaat   1320
tcaagtttaa cgaataaagg tggccatgcg gaatttgttt ctgattttgg ggcgaacaat   1380
gcgttcaaat acattaatgg ttcgtatgtg aaaaaacaag cggatggaaa attttactca   1440
ccggaagata ttgactatgg cacaggacct tctgtagtga aaatagtga ttggggacgt   1500
gtaggtcaca agaatgccta ctttggttca ggtgtaggtc tagccaatgg gcgtatttcc   1560
ttttcttttg gtatgacaac aaaaaggaaa agtaatgtgc ctgtatctag tgcgcaatgg   1620
tttgccttta gtactaactt aaatgcgcaa tcagtgaagc ctattttcaa ttatgggaat   1680
ccaaaagaac cagaaaaagc aacgattgaa ttcaatcgat acaaagccaa tgtcgttcct   1740
gtacttgtgc aaataaaga agtcactgat ggccagaaa atgtcaatga tttaaatgtg   1800
aagcgtgcg attctttaca atacattgtg acagggata cgacagaact tgccaaagta   1860
gatccaaaaa cagtaaccaa acaaggggatt cgagatacct ttgatgcaga aaaagtgacg   1920
attgatttat ccaaagtgaa agtttatcaa gcagacgcaa gtctaaacga gaaagactca   1980
aaagctgttg ctgcagcaat taattcagga agagctaaag acgtgactgc ttcttatgat   2040
cttcatttag atcaaaacac cgtcacagca atgatgaaaa ccaacgcaga cggttccatt   2100
gttttagcaa tggggtataa atatttactt gtcttgccgt ttgtagtgaa aaatgtggaa   2160
ggcgattttg aaaatacagc tgttcagctg acaaacgatg gtgaaacggt aacaaataca   2220
gtgattaacc atgtgccagg tagtaatcct tccaaagtg taaaagcaga taaaaacggt   2280
acagttggca gtgtttcgct acatgacaaa gatattccgt tacagacaaa aatttattat   2340
gaagtgaaat cttccgaacg tccagctaac tatggcggaa ttaccaaaga atggggcatg   2400
aatgatgtct ggacacgac ccatgatcgt ttcacaggga aatggcacgc tattacgaat   2460
tatgacctta agtaggggga taaaacgtta aaagcaggaa cagatatttc tgcctacatt   2520
ctttagaaa acaaagacaa taaagactta acgtttacga tgaatcaagc actattggcc   2580
gccttaaatg aaggaagcaa taaagtaggc aaacaagctt ggtctgtgta tctggaagtc   2640
gaacggatta aaacaggtga cgtagaaaac acgcaaacag aaaactacaa caagagctg   2700
atgcgttcta ataccgtggt gacgcatacg cctgatgatc caaaaccaac caaagccgtt   2760
cataacaaga aaggggaaga tattaaccat ggaaaagttg ctcgtggtga tgttctttct   2820
tatgaaatga cctgggactt aaaagggtac gataaagact ttgcctttga tacagtcgat   2880
cttgcgacag gcgtttcttt cttcgatgat tacgatgaaa cgaaggtgac accaatcaaa   2940
gacttacttc gtgtcaaaga ttctaaaggg gtagacatta cgaaccagtt cacgatcctca   3000
tgggacgatg ccaaaggcac ggtgacaatc tctgccaaag acccacaagc cttattcta   3060
gcgcatggtg ggcaagaatt gcgtgtaacg ctccctacaa aagtcaaagc cgatgtttct   3120
ggggatattt ataattcagc ggaacaaaat acatttggcc aacgaattaa aaccaatacc   3180
gttgtcaacc atattccaaa agtgaaccct aaaaaagacg tggttattaa agtcggtgac   3240
aaacaaagcc aaaatggcgc cacaatcaaa ttaggggaga aattcttcta tgaattcaca   3300
agtagtgaca ttcctgcaga atacgctggt gttgtggaag aatgtcgat tagcgataaa   3360
ctagacgtca aacatgacaa atttagtggc caatggtctg tgtttgccaa ttctaatttt   3420
gttttagcag acggaaccaa agtgaataaa ggggacgaca tttcgaaact attcacgatg   3480
accttgaac aaggggtagt gaaaatcacg gccagtcaag ccttttttgga tgcgatgaat   3540
ctaaaagaaa acaaaaacgt tgcgcactca tggaaagcgt tcattggtgt agaacgaatt   3600
```

```
gcggcaggag acgtttacaa cacgattgaa gaatctttca ataatgaaac gatcaagact   3660
aatacggtag tgacgcatac gccagaaaaa ccacaaacac caccgaaaaa aacagtgatc   3720
gtgccaccaa caccaaaaac accacaagca ccagtagagc cattagtggt agaaaaggca   3780
agtgtagagc cagaattgcc acaaacaggc gaaaaacaaa atgtcttatt aacggtagct   3840
ggaagtttag ctgcaatgct tggcttagca ggcttaggct ttaaacgtag aaaagaaaca   3900
aaa                                                                3903


Seq ID 12
atgaaaaaaa ttcttttttgc tagtttattt agtgccacac tactatttgg gggaagtgaa     60
atttctgctt ttgcacaaga aattatccct gatgatacta cgacaccgcc cattgaagta    120
ccaacagaac caagtacacc agaaaagcca acagatccaa caccgccaat tgagccacct    180
gtagaccctg tagagccacc tattacacca acggagccaa cagaaccgac agagccgaca    240
acaccaacag aacctacaac tcctacagag ccaagtgaac cagaacaacc aacggagcca    300
agtaaaccag tagaacctga aaaaccagtt acaccaagca aaccagcaga acccgaaaaa    360
actgtgacac caactaaacc aacagaatct gaaaaaccag tacaaccagc agaaccaagc    420
aagccaatcac acgttgttgt aacgccaaca ggggaattaa atcacgctgg aaatggtaca    480
caacagccaa cagtccctat tgaaacaagt aatttggcag aaatcacgca cgtgcctagt    540
gtaacaacac ctattacaac tacagacgga gaaaacattg tagctgtaga aaaaggtgtt    600
ccacttacac aaacagcaga agggttaaaa cctattcaat caagttacaa agtattgcct    660
agcggaaatg tagaagtaaa aggtaaggac ggtaaaatga aggttttacc atacacaggt    720
gaagaaatga atatcttttt atctgccgta ggcggtatct tgtctgtagt atctgggttt    780
gtcatctttta aaaacgcaa agctaaagta                                     810


Seq ID 13
atgggttccc ggttcggact gatgttgaaa caaggtggta acgatgtttt gctaatcgac     60
ggctggcagg agcatatcaa tgccatcaaa gaaaacggat tgaaggctaa ttataatggc    120
gaagaaatta ctgttaaagt ctcaattgtt aatcagaatg aggtgcccac tggtgagcaa    180
tttgatttga ttatcttatt cacgaaggca atgcagttgg aaaaaatgct gcaggatgtt    240
aaaccattaa ttgctgatca cacagaagtt ttgtgtcttt taaacggtat cggtcatgaa    300
gacgtaatcg aaaaattcgt accgatggaa aaaatctttta tcggcaacac catgtggact    360
gctggtctag aaggtcccgg caaggctaaa ttatttggca gcgggtctgt tgaattgcaa    420
aatctaggta ttggccaaga agaatcagcg aaaaaattgg cggaaacttt gtccgcgtct    480
ggtttaaatg ccaaatattc cgacaacatt cattattcaa tttatcgtaa agcttgtgtc    540
aatggtacga tgaatggctt atgcacgatt ttagacgtca atatggcggg gcttggtgca    600
accaaacctg ctcatgacat ggtagtgact attgttaacg aatttgcagc tgttgctgcc    660
aaagaaacg tcaacttaga cattcccgaa gtaatcgaac atgtagaaac ttgttttgac    720
ccgactacga ttggtatgca tttcccatcg atgcaccaag acttgataaa aaataaccgt    780
ttgaccgaga ttgactacat caacggagcc atttctcgta aaggtaaaaa atatggagta    840
gtaactcctt actgtgattt cttaacgcaa ctggttcata gtaaagaaga aattctagga    900
gcgaaa                                                              906


Seq ID 14
atgcagacaa cagaagaaat gcaatctggc gatggacgcc gaggtcgatt gcggatattt     60
tttggatttg cggctggtgt tgggaaaacc tacggtatgt tgacagaggc tcatgaatta    120
ttgatgatgg gaaaacatgt tgttgtcggc tatatcgaac ctcatgatcg tccagatacc    180
aatcgattat tagaaggatt accgcaaatt cctcctaaga atatttttta taaacaaatg    240
gttttaacag agccggatat tgatcaaatc attcaacaga agcccgaaat cgtgttgatt    300
gatgaactgg ctcattccaa tgcggaaggt tctcgaaatc gtaaacgata ccaggatatt    360
gatgaactgt tgaatgcagg aatcgatgtt tttactacgg tcaatgttca acatatcgaa    420
agtttgaatg atatcgttga agaagttacc ggaattgagg tcaaggaaac ggtgccggac    480
acttttctac ggcaagcgac aataagagtg attgatgtgg agcccgatga gctgatcgaa    540
cgacttgaac aaggaaaaat atatgcgaat gaaaacgcga acgtgccttg aaaaatttt    600
tttattccaa aaaattgga ccaattacgt ggattagcaa ttcagcgagc gtcggatcat    660
attaatcgaa ttagtggaaa gacaattggg atacaaagta aattattgac ggtggtgaac    720
gatgcctttc caaaatgac cgaaaaatgt attcgctgga cagcacgttt agcgcaaggt    780
ctagttgtcg attggacagt gattcaagtt cgtacgcaag aaatacacc aacgaacatt    840
cctttagccg ataaaattagg tgcggaggtt attagtattg aagaggatga tagtttttgag    900
acgattgtcg agttcgcaaa aatgacgggt gttactgata ttattatggg gaaaaatcta    960
aggcagccgt ggtatgaaaa aatcttttat gaagcgtttg atgatcgctt gttaaaacga   1020
ttgaaagata cggaacttca tttaattcct ttcaatgaag aaaagcggtc tttatttttt   1080
aagactcgta aagtcattga gggtggcgga aagagattag tgattgcaat tggtggtgtt   1140
tttctagcga cggtcgtgac ggaattgatg caatatatcc acgttggcga tcaaaatctg   1200
atgcttatt atatttcatt cgtattatta gtcgctcgaa cgacatccgg ttatttttgg   1260
agctcattat cctctatttt aagtgttttg tcttttaact ggttttttgt cgaaccatta   1320
tattctctga cagtatataa acaggatat cctttttactt tattgctcat gctagttgtt   1380
gccctgatga gcagtaattt aatgattcga ttaaagaaac aagcagatac ttctatgaaa   1440
aaagacacc aaatggaaat tttgtacgaa ttgaacaaac gatatgtttt agtggaaagt   1500
cgaaaacaaa ttcttgatat ttctgccact tatttatcaa gattattgga acgagaagtg   1560
attattttttg atcgtcaagt gaagacagag agtgttcatt gtatcaacga aaaaaaatca   1620
atattaaata tgaagacga agcagctgta gccttttggg cagcaaagaa ccaaaaagaa   1680
gctgggaatg gtacagacac actaaatgga gctaaaggat tctatttgcc gattgctgca   1740
```

```
ggcaggaaga cgttggcggt attaggcatt gaacgaaatg cagatcttga tttggaaaat   1800
gatcagctca attacttaaa attggtgctg actcaaattg cggtcatctt ggaacaaaca   1860
gaattgaaag atgagaaaga gcaagttgag ttagaaaatg aacgggaaaa agtccgcagc   1920
aatttattac gagccgtatc tcatgatctg agaacaccat taactgttat ctcagggatt   1980
gctgaaacat taggcattgg aaatgattta aaagaagaga cacagcgaaa attattaaag   2040
gatattcaag aagagtcaca atggctgatt cggatggttg aaaatctttt atccatcacg   2100
cgaatcaaca tggatactat gaaggttaat aaaactgctg aaccagtaga agaggtcatc   2160
gaggctgttt ataagcattt gaggaaggtc tatccagacg gccaagtgga tattcacctg   2220
ccagaagagg tcattttat tcaagcggac cccattttga ttgaacaagc gttgtttaat   2280
ctgatagaaa atgcgtttcg tcatggagaa aatgatctac cagtcaaact aaacgtctat   2340
caggaaaaag aacaaacagt ttttgaaatt gaaaatcacg gcgaaattcc attaaaacaa   2400
tttcaaaaga tccaatcaaa cttgtccggc acaaatgaag taccggttga ctcaaaaaat   2460
ggtttaggaa tcggattgag tattgtaaaa acaatcgtcc atgcgcacaa tggtaaaatg   2520
gaaatgacga taggtagggg caaaacattg gttcgaattt atttgaag             2568


Seq ID 15
atgaaaaatt tttataaaaa aaaatttgct ttaactgatc agggcgcaga agctttaact    60
aaagcttcta tttctagttt ttttgtatac tgcattaata tggtaccagc attcattata   120
atgatgctga tagacgaatt agttttagaa aatgcaaagc cccgttggct ttattttgca   180
gtatcgtttg ttacttttgct tttcatgtat tggttgttag ataggagta tgaaaactta   240
tataatagta cttataaaga aagtgcgcat ttaagagtgc aaattgcaga cgatttgtca   300
aatttaccat tatcctattt ttcaaaacat aatttatcag atttatctca aactatcatg   360
tctgacgttg aaggtattga gcatgcgatg agtcatgcaa tacctaaatc cggcggtatg   420
gctctgtttt tcccttttat ttcagtgatg ctttggttg gtaatgtcaa aatgggatta   480
gctgttattt tgccaacgtt atttagtttt gtcttaatct tgttatcaaa gaaatcccaa   540
acgaaagcca atactaaata ttacgatact ttgagagaaa actcggaaga atttcaagaa   600
actattgaac tgcagcaaga gattaatagc tttaatctat ctaaaaagt tcaagacaga   660
cttttcaaaa aaatggaaga gagtgaaagg attcatttaa aggtagaatt aagtactttt   720
tcagtcatgg ccttatcctc tattttctca tatgttagtt tagcagtagt cattctagta   780
ggtgttcact tactgttaac gggtgaagtg actatactct acgtcgttgg ttacttacta   840
gccgcaataa aaataaagga ttccttcgat tctatgaaag aagcagtgct tgaaatattt   900
tatttagctc ctaaaataca gcgaattaga gctatgaagg aaaccagtat ccaagaggga   960
tcagatagtc ctctgaaatc ttttgacgtt gagttgcgag atgtctcttt ctcatatgat  1020
aacaataccc caattttaga tcatatctct tttacggcta aacaaggaga agtgactgcg  1080
ttagtaggtg ctagtggttc cgggaaaact agtatcttga aattagtgtc tagactgtat  1140
gattatgatg aaggctgtat tttaattgat gggtatgata ttaaacgtgt atctccagct  1200
tctttgttta gcaaaatcgc tattgtttta caagaggtaa ctttgtttaa tacaagtatt  1260
ttggaaaata tacgtattgg aaatagtcaa gcgtcagatg aggaagtgaa aaaagcagct  1320
cgtttagcta actgcgaaga ctttattgaa aaactacctg acggttatca tacattggtc  1380
ggagaaaacg gaagttcact ttctggagga gagcgacaac gcttatccat agcacgggct  1440
tttttgaaaa atgcgccaat tttaatttta gatgaaatta ctgccagttt agatgcggaa  1500
aatgaaaaac gtattcagga aagtttaaat cgtttaatac aagataagac agtactaatt  1560
atatcgcatc gcttgaagtc tattgaaaag gtaaataaaa ttgtagtgat ggatcaagga  1620
aaagtagttg atcaaggtac tcatagtgag ttgtatcgaa ggtcagaaat atacaaaaat  1680
ctcataaaga aaactaaact atctgagaaa ttcgtctatg aaaaagaagc tcaatcacga  1740


Seq ID 16
ttgactcttt ccgacacgtt aggtcaacga gaagtaacgg ctctttctta cgattcgaga    60
gacgtaacag cagaaacact cttttttgt aaaggattaa acttaaaga aatctattta   120
gaaaatgctg taaaagacgg cctagaaatc tatgtttctg aggtgcctta tgaggttccg   180
gcgcaactag gaattatcgt gactgacatt aaaaaagcaa tggctgtgtt aagcatggct   240
ttttatgatt atccacaaaa taaattaaaa ctgattggct ttactggaac caagggtaaa   300
acaacagcgg cttatttcac caaatatatt ttagatgttg ctacacaaca aaaaacagcc   360
ttgctttcta ccatgaattc gacattagat ggcaaaacat tctttaaatc tgccttgact   420
acacccgaat ctttggactt gtatcgaatg atggcaaccg ctgttgctaa cgggatgact   480
cattttatta tggaagtttc ttcccaagcc tacaaaacta atcgtgtcta caagttattt   540
tttgatgtcg gcattttttt aaacattacc cccgaccata tcagtccaat tgaacatcca   600
acctttgacg attacttta ttgcaaacgt caattgatta cacattctaa agtcattgtg   660
ttgaatcacg aagctgatta ttttccatta cttaaagaaa ctgcacaaca gcaaaaagtc   720
ccagccattg tctatggcag tcagccggct ccagaagtag actattcatt cgctgtttct   780
tcagaggatt cccttacgtt tatcgttgaa agtcctgcag atgcgttagg attggccggc   840
tcttatcacc ttcgtttagg tggtgatttt aacaaaggta atgcggttaag tgcggcgatt   900
gctagtgtcc ttgtgggagc ttcaaaagaa gagtgtcaac aaggaatcgc agcaaccacc   960
gtccctggcc ggatggaatc tttaacgaat accaacggtg caaccgtcta tgttgattac  1020
gctcataact atgacagttt gaagaatctt ttaacctttg tccgtgaaga acatcctgac  1080
ggccgcttaa ttgttcttgt cggtagcacg ggagacaaag cgatttctcg ccgcaaagac  1140
tttggtaggg ttctttctga attggctgat gttgcagtcc taacaaccga tgacccagcc  1200
agtgaagatc cagcaaaaat ttgccaagaa attcaagccc atattaccaa agaaatgcct  1260
gtttacaccg ttttagatcg tggcgaagca attgctcacg cattatcttt aagcacaacg  1320
gctgatgatg cgattgtctt agctggtaaa ggcgccgacc tttatcaaaa agtaaacggc  1380
gtagatgaac cttacgcagg agactttgcc cttgcggaag ctttcattaa taaaaagaac  1440
```

Seq ID 17
```
atgtttttag gatttaatga aatgaagtat tcaaaaggac ggtacgtcct tgtagtctta   60
gtaatggtct tgattgcttg gttaattttt attttgtcag gtttagcgaa cggtttagct  120
caaggaaatc gtttagccgt agatcaatgg caagcgaatc aagtggtttt atcaaaagag  180
gccaatagta atttaaatgt atcagtgtta gatgaaacg tgaaagaaac gatttcagga   240
ggcaaaattg caccgattgg tcaacaatct ttagccatcc gcccagcaga tgataaaaag  300
gctgaattaa cgaatgttag cttatttggg attgaaaaag aaagtttttt aatgccaaaa  360
gtgattgaag gaaacgcatt tactgataaa aatcaagtga ttgcttcaga aacattgaaa  420
aatcaaggat ttaaaattgg cgacaaatta actgcaggga aatatgatga gcaattagaa  480
attgtcggct ttatttctaa aagtagctat aacattgtgc cagttattta cacttctctg  540
gatacttggc ggtcaattaa atatggtgac aatccagcaa tggccaaaat ggttaatggt  600
ttcattgttc gcagcaagga caacacggaa gttaaaacga ctaataaaga cagccaagtt  660
ctttcaattt cagattttat tgaaaaactg ccaggataca gcgctcaaaa cttgaccttg  720
gatggcatga tttatttctt gattgtgatt gcagcgttca ttatcggtat ttttatcttt  780
gtgatgactt tacagaaaac agcaatgttc ggcgtcttaa aagttcaagg ggtgccaact  840
agtttcttag caaaggcagt catgctacaa accgctttat tggctgttct aggagtggcg  900
attggtcttg ctttaacggg aattaccgtg ctcttttac cagaagcgat gcctatgca   960
actaatggac cacgcatgat tttgtttagt gtactattga ttttatctgc attaattggc 1020
ggggcatttt caattcgaac gattgctaaa attgatccgt taatcgcgat tggaggt     1077
```

Seq ID 18
```
atgcaaattt ttccctattc agatgatcca aataaacgct atcacacttg gaactatgca   60
cttcgtcaac agtttggtga aaagattttt aaagtgccaa ttgatggtgg ctttgattgc  120
cccaatcgtg atgggacggt tgccaaaggt ggttgtacgt tctgtagtgt ttctggctca  180
ggcgacatga ttgttgcacc tagcgatccc ctgccgcttc aatttcaaaa agaaatccaa  240
ttgatgcatc aaaaatggcc aactgtcgat caatacattg ttacttcca aaatttcacc   300
aatacacatg cgcccgttga cgttattcgt catcggtttg aacaagtagt caatgaaaaa  360
ggtgtcgttg gtttatccat tggtacccgt cctgattgtt tgcctgatga ggtagtgaac  420
tatttagccg aattaaatga acgcttttat ttatgggtag agttggggttt gcaaaccacg  480
tttgaggaaa ctagcgctgc cattaatcgt gctcatgatt atcaaacgt tctataactg  540
gtggctaaat tacggaaaca cgggattcgg gtctgtacac atttaattaa cgggctacct  600
ggggaaacac ctgccatgat gcgcgaaaat gtgcggcgaa ctattcagga ttcagatatc  660
caagggatta aactgcactt acttcattta atgacaaata ccaaaatgat gcgggattat  720
aatgaaggac gcttacaatt gatgagcaaa gaagcctatg tttcagtgat ttgtgatcag  780
ttagaaatga ttccacccga aattgttatt catcgcttga caggcgatgc tcccttgaa   840
acaattatcg gaccgatgtg gagcttgaaa aaatgggaag ttctcaatgc gattgatgcg  900
gaaatgaaac gccgcaacag ttatcaagga aaatatactg tcatttcagg aaaggaagtt  960
ttcaat                                                                966
```

Seq ID 19
```
atgaagaaag aatcaatgtc acgtatcgaa agaaggaaag cacaacaaag aaagaaaacg   60
ccagtacaat ggaagaagag cactacttta ttcgctcgg cgttaattgt ttcatctgta  120
ggaacgcccg ttgcgttact accagtgact gctgaggcaa cagaagagca gccaacaaat  180
gcggaagttg cccaagcacc tactacgaa actggcttag tagagacacc aacaacagaa   240
actacgccag gaactacgga acaaccgaca acggattcgt caacaacgac tgaatcgaca  300
actgaatcat caaaagaaac accaacaaca ccaagtaccg agcaaccaac agctgattca  360
actacacctg tggaatcagg aacgactgat tcttcagtag cagaaattac gccagtagct  420
ccttcagcaa ccgagtctga agcagcgcct gcggttacac ccgatgatga agtaaaagta  480
ccagaagcta gagtagcttc tgcgcaaact ttttcagcgt tatcaccgac gcaaagtcct  540
tcagaattta ttgccgagtt agctcgttgt gcacaaccta ttgcgcaagc caatgattta  600
tatgcatcag tgatgatggc tcaagcaatc gttgaaagtg gttggggagc aagtacgcta  660
tctaaggcac caaactataa cttatttggg attaaaggca gctacaatgg acaatctgtc  720
tatatggata catgggaata tttaaacggc aaatggttag tgaaaaaaga acctttccgt  780
aaatatcctt cttacatgga atcattccaa gataatggcgc acgtgctaaa aacaacttct  840
ttccaagcgg gcgtttacta ttatgctggg gcttggaaaa gcaatacaag ctcgtaccgc  900
gatgcaactg cttggttaac aggtcgttat gcgacagatc ctagctacaa tgctaaatta  960
aataatgtca ttaccgcata taacttaact caatatgata caccatcttc tggtggaaat 1020
actgggggcg gaacagttaa tccaggaaca ggcggctcga caatcaatc aggaacgaac 1080
acgtactata ctgtaaaatc aggagatacc ttgaataaaa ttgccgcgca atatggtgtg 1140
agcgttgcta atttacgctc atggaacggc atctctggcg atttaatttt cgttggtcaa 1200
aaacttatcg tgaaaaaagg tgcttcaggt aacactggtg gctcaggcag cggtggttct 1260
aacaataatc aatcaggaac gaacacgtac tatactgtaa aatcagggga taccttgaat 1320
aaaattgccg cccaatatgg cgtgagcgtt gctaatttac gctcatggaa tggcatctct 1380
ggcgatttaa ttttcgttgg tcaaaaactc atcgtgaaaa aaggtgcttc aggtaacact 1440
ggtggctcaa caacggtgg ctctaacaat aatcaatcag gaacgaatac gtactacaca 1500
attaaatcag gcgataccct gaacaaaatt gccgcccaat atggcgtgag tgttgctaat 1560
ttacgctcat ggaatggcat ttctggcgat ttaatcttcg ctggtcaaaa aattattgtg 1620
aaaaaaggta cttcaggtaa caccggtggc tcaagcaatg gtggttctaa caataatcaa 1680
tcaggaacga atacgtacta cacgattaaa tcgggcgata ccttgaacaa aaatttctgca 1740
caattcggtg ttagtgtggc taacttacaa gcctggaata acatcagcgg tagtttgatt 1800
```

```
tttgctggtc aaaagattat cgtgaaaaaa ggcgccaact caggttcaac gaatacgaac   1860
aagcctacga ataatggtgg cggtgcgaca acatcctaca cgattaaatc aggtgatacg   1920
ctgaataaaa tttctgcaca gtttggcgtg agtgttgcta atctacgttc atggaacggg   1980
atcaaaggcg atttaatttt tgctggtcaa acaatcatcg tgaaaaaagg cgcttctgca   2040
ggtggcaatg cttcttcaac aaatagtgca tcaggcaaac gccatacagt taaaagcggt   2100
gattcacttt ggggcttatc aatgcaaatc ggaatcagca tccaaaaaat caaacaatta   2160
aatggcttaa gcgggggatac aatttatatt ggtcaaactt taaaagttgg t           2211
```

Seq ID 20
```
atgatgattg ttgtactttt aatgagtagt tgcttatttt tatttgtaag ctatcaattg     60
tatttgcgat tacttgtcga aaaactggat tcagaggacc ggacattaaa gaggaaaata    120
aatttctata agtatcagaa aaatagtcgt ttaatgattt atttattgtt tagtgtttgt    180
ctacttagcg tattattgtt aggaattgtt tatacatatt atcaattaaa tcaacgaaat    240
atccgcatgg aacaacggat cgagcgtttg gcacagggcc aagctactca aggtgacaaa    300
ataaaaaaac aagcgattaa aaaaacagcc ttaaatcagt ttccctggaa gcaagctgtt    360
tcggcagaat cggccgctgt cttgactaat tatgaattac aattagcgcg tgaatggcgt    420
ccttacctcg gagacgacag tatcacaatg attagaagtg aaaaaacaca aaccttgaca    480
ctgtctgttt tttcagttgg cttaagttac catgaatttc aaacaggcca agacaatatt    540
gtagcactta ttgcagcgtt gaattcagtt aaagaaatta ccatgattga ttttaatttt    600
acttatcgag atcaagagca aacccttgtc aaatccattg atacgtatgc tcgtgagacg    660
ttggaaacaa atttggagcc tgtcgtgatt agt                                  693
```

Seq ID 21
```
atgtctaact atactattat cgatttaaac caacaaccac acttaattcc gtcagctgct     60
caatggttta gccaaaagtg gcaaatcccc gctgaagcgt atcaagcaag tatgcaggaa    120
atgctaacta gcgaaaacca cgtaccacat tggtgggtct gcttgaataa acagcaggac    180
atcgttgcag gcgtgggcgt aattgataat gacttccaca atcgtgcaga ccttacgcca    240
aacctttgtg cattgtatgt cgaaccaaca tttcgtcacc aaggcttagc aggaattctt    300
ttaacaacag ttggtgattt tctagctaat tctggatttа aaaagcttta tttgctcaca    360
gaccatacta cttttatga aagatatgac tgggaatttt tgacaatggt aaccaccgaa     420
gagcagtctt cgctagaat ttatcaaaaa caattacgtc aa                        462
```

Seq ID 22
```
atgaaaaatt cagaaaatga ttatatccaa tctttattcc aaattcttcc cggcttgcta     60
acagcttcc ttgtcgcctg ccttagcaag tttcttgcca tctggttacc tagcttaggc     120
gccgctacga tagctatttt actaggaatc tttttgggaa atacctttgt tcgcctgtcc    180
aatttaaatc gcggaaccaa agtagccgaa agtaaattac tagaattttc tgttgtttta    240
ttgggaacca ccgtgacctt tcaaacgatt gcccaaattg gcttacaagg tgtcgcattt    300
attctcattc aaatgtcgct gaccattatt tttgcctatt tgatcggtaa aaaactggct    360
ttttctgata atatgtcttt attgatggct ggtggtaatg ccgtttgtgg gtcttctgct    420
attgcttcga ttgctccagc aattcaagca gatgaggaag aaaaaggcca aatcattact    480
ttagtaaact tattaggaac cgttttaatg ttgaccttac ccattttaag tgggattctc    540
tatggtacca atttgttggc tcgtagcgca ctaatcggag ggactttaca atcagttggt    600
caggttgtcg ccagtgccaa tatggtcaat gagaacgcag ttcaacttgc gatgctcttt    660
aaaattatgc ggattgtcct actcgtagca gttgtctatt tatttggacg tttcaagcaa    720
agtaagacga cagaatcaga ggctgagttg gtagaagtca ccaaaaaaag cagcgcccta    780
ccttggtatg tagttggctt tttcattgcc tgtgtcttta atagtttgat tcatttcccc    840
gtcgtgatca gtgagactgc tcattctttt agttcttggt ttgaaattac tgccttggca    900
gcaatcgggt tacgactcga ttttaaaaag tttttccaag aaggcaaacg tttcttaatc    960
tatggtttat cagtaggaac tgttcaagtc gtcctagcca ttttgttgtt ggcactgcta   1020
cagttc                                                              1026
```

Seq ID 23
```
atggcaaaaa caggaatgta tgtaggcctc gatattggta caacttctgt caaagtagtt     60
gtcgctgaat atattgaagg tcaaatgaac atcatcggcg ttggaaatgc aaaatcagac    120
ggtttgaatc gagggattgt cgtcgatatt gatcaaacag ttcaagcaat ccaacgggcc    180
gttcgacaag ccgaagaaaa agccggcatt caaattaaaa gcgtaaatgt tggattacca    240
gcgaattat tagaagtaga aagctgtcaa ggaatgatag ctgttagcag tgaatcaaaa    300
gaaattacag atgaggatgt gcgcaacgtc gcatctgcag ctttggttcg ttcaactcct    360
cctgaacgtc aaattgttgc aatcttaccg caagactta ccgtcgacgg ctttgaaggc     420
atcaaagacc cacgtggaat gcttggcgtt cgcatggaca tgtttggcgt tgtctatact    480
ggaccaaaaa caatcatcca taacatccgt aaatgtgttg aaaaagctgg cttaggaatt    540
aacgaattag tgattacacc gcttgcttta acggaaacaa ttttaacaga tggcgaaaaa    600
gactttggaa caattgtcat tgatatgggt ggcggtcaaa caacgacatc tgttatccac    660
gacaaacaat tgaaatttac tcatgttaac caagaaggcg gcgagttcat cacgaaagat    720
atctcaatcg tattgaatac ttctttttaac aacgccgaag ccttaaaaat taattatggg    780
gatgcttatc cagaaagaac ttcagcaaac gaagagtttc tcgttgatgt tatcggtaag    840
tcagagcctg taagagtgga cgaacgttac ctttcagaaa ttattgaagc gcgtgtggaa    900
caaatttga gaaatcaaa agaagttctg atgaaattg atgcgtttga attacctgga      960
ggcgttgtgt aactggtgg cgccgctagt atgccaggta ttgtcgattt ggcgcaagag   1020
atttttgaag ccaacgtaaa actatacgta ccaaatcaca tgggcttacg taacccagtc   1080
```

```
tttgctaatg tcattagcat tgtggaatat tcagcgcaac taaacgatat ttatcacatt   1140
gccaaatatg caatccctgg tgaaaaatcg aaaccagcac aatcagtcgc tgtccaacaa   1200
gaagttcgct atgacacata tgctgaacag cctcaagaag aatacgaaga attcaacgaa   1260
cgtgagtctg gtgaaaaagt gacaggcaaa atcaaagact tcttctcgaa cattttcgac   1320
```

Seq ID 24

```
atggcaaaaa aaacaattat gttagtttgt tccgcaggaa tgagcacgag tttattagta    60
acaaaaatgc aaaaagcagc agaagatcgt ggcatggaag cagacatctt tgcagtatcg   120
gcttctgaag cagatacaaa cttggaaaat aaagaggtga atgtttact tttaggtcca    180
caagttcgtt tcatgaaagg gcaatttgaa caaaaattac aaccaaaagg gattcctta    240
gatgtaatta acatggcaga ttatggcatg atgaatggcg aaaaagtttt agatcaagca   300
atctcattaa tggga                                                    315
```

Seq ID 25

```
atgaaaaaag agatcattgt ttatactatt tccgattcac ttggagaaac atcacaaaaa    60
ttgttagctg cagcaagcgc acaatatcca gatatttctt ttcttaatcg atacaatttt   120
tcttttgtca caacagaaga agaattatta gagattttaa aagatgcctt aaaagataaa   180
gccttagttg tcagtacatt agtcagtaaa caactaatca cagcggcgaa agaatttagt   240
gaacgaacag ggttgttata tttagattta atggcgccat tttttgaatt aattcaagcg   300
aaagccggag tagatcctat tgaagagcct ggacgacgcc accaactaga tcgtgcctat   360
tttgataaaa tctcagcgat tgaatttgct gtaaaatatg atgatggcaa aaatcctcaa   420
gggtttcttg attctgatat attgttgtta ggcgtttcgc ggacctcaaa gacgccagtc   480
agtatgtatt tagcgaatca aggctaccgc gtttctaact taccattaat tccagaagtt   540
ccattgccgc caatttttgga agaaatggat ccacaaaaaa tgattggttt agtttgttcg   600
ccagaaacat taggacagat tcgtagcagt cggttggctt ccttaggttt aggtaatgag   660
accagttata ccaatgttga acggattgaa caagaattag cttatgccga agagattttt   720
gcgaagtatg gcatcccagt gattgatgta accgcaaaat ctgtcgaaga aacagccttt   780
ttaattaaag aaaaactaga tgaaagaaat                                    810
```

Seq ID 26

```
atggaatgga ttgaaattaa gcacgcaaca caaaataatt tgaagaatat ctctgtcaat    60
atccctaaaa agcaactaac tgttgttact ggactttcgg gttcaggaaa gtcctcctta   120
gtatttgaca cattagccgc tgaatcacgt cgtgaactaa atgatacatt tagttcgttt   180
gttcaaaatt acttacccaa atatggtcgt ccagaagttg aaaaaatcga gaatcttcct   240
gttgcaattg tcattgacca gaaaaaagta gcaggtaatt ctcgttcgac agtgggaacg   300
tacacagata tttatacgtt tttacgtttg ctgtttcac gagcaggttc tccatttgtt   360
ggttattcag atactttttc gtttaatcat ccggatggaa aatgtccaac atgcgatggc   420
ttaggaaaga ttaccgaaat caatcttcat cagctagttg attatgacaa atcattaaac   480
aaagggccta tcgatttccc cacatttaca gttggcaact ggcggtggaa gcgttatgcc   540
catagtggtc tatttgattt ggataaaaaa atcaaggact attctcctga agagttagca   600
ttattttat atgctccaca acaaaaacta gctaatccac ccaaagagtg gcctcataca   660
gctttgtatg aaggaatcgt cccgcgtatg caacgtagca tattgcatac agacgaaggc   720
aaacgtcatc aaaaatacct taatcacttt gttaccgtaa aaagatgtcc tgattgttta   780
ggaagtagag tcaatgaacg tgttcgtagc tgcaaaatta atcagaaaag tattgctgat   840
gctgttgaca tgccactcac tgaattacat tcttttattc gttcaatgga cctatcctta   900
ataaaaacta ttcaagaaga gctacttgta cgtctagaag cattaattaa tatcggtctc   960
tcctacctca cattaggacg agcaactgaa acactttctg gtggtgaagc acagcggatt  1020
aaaattgcta agtatgtaaa cagcgcctta aatgatatta tgtatatttt agatgaacca  1080
agtgctggct tacatccaaa ggacatcgaa cggatcagtc gtgcattgct caatttaaaa  1140
aataaaggaa acaccgtggt tctcgtggaa cataatccac aattaattag agaagctgat  1200
tttatcatcg atatcggacc tttcgcaggc gaaaatggtg gccatgtcca gttttcagga  1260
acgtatgacg cattttttagt ctccaaaacc ttgactagtc aagcgcttca agagccgtc   1320
cctttaaacg accaaccaag aaaagcaagg aagtctttat caatcgaaca tgcgacactg  1380
cataatttaa acaatctatc tgtcgaagtt ccgttaggag tttttgactgt tatttgtggg  1440
gtcgctggtt caggtaaatc atcattagcg gaagaaattt atcaaaaagc ccaggctgat  1500
aaccaagaaa ttattcatct ttcacaaaaa agcattaccg caaatttacg atccacacct  1560
atgacctatc ttaatatttt tgataaggtc cgcaaactgt ttgcggaaga aaatcatgtt  1620
agtccagctt tatttagcta taattccaaa ggcgcctgtc ctacttgcaa ggggaaaggc  1680
ataattgtct ccgatatgtc ttttatggag gatgttacta gtatttgcga aacctgtcac  1740
ggaacgcgtt acaaaagaga ggtgcttcat tatctgtata acggaaaaa tatcgttgaa  1800
gtactagctt taagcgttaa agatggctat gattttttca aagaccagcc ttttgctctt  1860
tcattaaaaa atttactgga agttggctta agctatctta aactcaatca atcgctctca  1920
acattatctg gtggcgaatt gcaaagggta aaattagcag acacacttca ccaaaaaaaa  1980
gctatctatt taatggatga acctacagat ggtttacact taattgatat ccaacaaagt  2040
cttcaacttt tcaatcgaat ggtagaagaa ggaaacagct tgatttttatt agaacatcac  2100
attgatgtga ttaaaagtgc cgactggttg attgaactag tcccgaagg tggaaaaaat  2160
ggcggacagc ttcttttttac aggaacccca gcaaatatgc taaattccac tcattctatt  2220
actaaaggct atctg                                                    2235
```

Seq ID 27

```
atgaagaaat taaaaatgtt aggatgcgtc gggttgcttt tagctttaac ggcttgtcag    60
```

```
gcgggaacgg gaaactcggc tgatagtaac aaagcagcg aacaaaaaat tgcaattagt    120
tctgaagcgg ctatttcgac aatggaacca cacacagcgg gggatacgac ctcgacttta    180
gtcatgaatc aagtttatga aggactctat gttttaggta aagaagatga attagagttg    240
ggggtcgctg ccgaagaacc agcgattct gaagatgaaa ccgtttatac atttaagatt    300
agagaagatg ccaaatggtc gaatgatgat ccagtaacaa caaacgactt tgtttatgca    360
tggcaacaag ttgcttcccc taaatcagga tcgattcatc aagctttatt ttttgatgtc    420
attaaaaatg ctaaggaaat tgctttagaa ggcgcagatg tgaatactct tggggttaag    480
gcgctagatg ataaaacgtt agaaataact ttagaacggc ccacccctta tttgaaatca    540
ttactttcgt ttcctgtttt gtttccacaa aatgaaaaat atatcaaaga acaaggggat    600
aaatatgcta ctgatgcaga acatttgatt tataatggtc cttttaaatt gaaagaatgg    660
gataatgcct cttctgatga ctggacctac gaaaaaaatg atacgtattg ggatgctgaa    720
aaagttaaat taacagaagc gaaagtttca gtaattaaga gcccaacgac agcggtgaat    780
ttgtttgact cgaatgaatt ggatgtagtg aataagctaa gtggtgaatt tattcctggt    840
tatgttgata atccagcctt tctttcaatt cctcaattcg tcacatactt tttaaaaatg    900
aacagcgttc gtgatggaaa agaaaatccg gctttagcga caacaatat tcgtaaagcg    960
ttggcacaag cttttgataa agaaagtttt gtaaaagaag tcttgcaaga tcaatcaacg    1020
gctacagatc aagtaattcc gccgggacaa acgattgcgc cagatggaac agatttcaca    1080
aaactagctg ctaagaaaaa taactactta acctacgata cagcgaaagc aaaagaattc    1140
tgggaaaaag ggaaaaaaga aattgggctg gataaaatca aattagaatt tttaacagat    1200
gatacagaca gcgccaaaaa agctgctgag tttttccaat ttcaattgga agaaaatcta    1260
gatggattag aagtgaatgt tactcaagtt cctttacta ttcgtgttga tcgtgatcaa    1320
acgagagact atgatttaga attatctggt tggggaaccg attatcgtga tccattaaca    1380
gttatgcgca tcttttacttc ggatagtacc ttgggcggcg taacgttcaa gagtgatacg    1440
tatgatcaat taattcaaga aactagaaca acacatgcgg ctgatcaaga ggctcgttta    1500
aatgactttg ctcaagcaca agatattttg gtgaatcagg aaacggtttt agcaccaatc    1560
tacaatcgaa gcatttctgt attagctaat caaaaaatca aggatctgta ttggcattca    1620
tttggaccca cgtacagttt aaaatgggct tatgttaac                           1659
```

Seq ID 28
```
atgaagcaat taaaaaaagt ttggtacacc gttagtacct tgttactaat tttgccactt     60
ttcacaagtg tattagggac aacaactgca ttgcagaag aaaatgggga gagcgcacag    120
ctcgtgattc acaaaaagaa aatgacggat ttaccagatc cgcttattca aaatagcggg    180
aaagaaatga gcgagtttga taaatatcaa ggactggcag atgtgacgtt tagtatttat    240
aacgtgacga acgaatttta cgagcaacga gcggcaggcg caagcgttga tgcagctaaa    300
caagctgtcc aaagtttaac tcctgggaaa cctgttgctc aaggaaccac cgatgcaaat    360
gggaatggtca ctgttcagtt acctaaaaaa caaaatggt aagatgcagt gtataccatt    420
aaagaagaac caaaagaggg tgtagttgct gctacgaata tggtggtggc gttcccagtt    480
tacgaaatga tcaagcaaac agatggttcc tataaatatg gaacgaaga attagcggtt    540
gttcatattt atcctaaaaa tgtggtagcc aatgatggta gtttgaaaaa agtaggaact    600
gctgaaaatg aaggattaaa tggcgcagaa tttgttattt ctaaaagcga aggctcacca    660
ggcacagtaa aatatatcca aggagtcaaa gatggattat atacatggac aacggataaa    720
gaacaagcaa aacgctttat tactgggaaa agttatgaaa ttggcgaaa tgatttcaca    780
gaagcagaga atggaacggg agaattaaca gttaaaaatc ttgaggttgg ttcgtatatt    840
ttagaagaag taaaagctcc aaataatgca gaattaattg aaaatcaaac aaaaacacca    900
tttacaattg aagcaaacaa tcaaacacct gttgaaaaaa cagtcaaaaa tgatacctct    960
aaagttgata aaacaacacc aagcttagat ggtaaagatg tggcaattgg cgaaaaaatt    1020
aaatatcaaa tttctgtaaa tattccattg gggattgcag acaaagaagg cgacgctaat    1080
aaatacgtca aattcaattt agttgataaa catgatgcag ccttaacttt tgataacgtg    1140
acttctggag agtatgctta tgcgttatat gatgggggata cagtgattgc tcctgaaaat    1200
tatcaagtga ctgaacaagc aaatggcttc actgtcgccg ttaatccagc gtatattcct    1260
acgctaacac caggcggcac actaaaattc gtttactta tgcatttaaa tgaaaaagca    1320
gatcctacga aaggctttaa aaatgaggcg aatgttgata acggtcatac cgacgaccaa    1380
acaccaccaa ctgttgaagt tgtgacaggt gggaaacgtt tcattaaagt cgatggcgat    1440
gtgacagcga cacaagcctt ggcgggagct tcctttgtcg tccgtgatca aaacagcgac    1500
acagcaaatt atttggaaat cgatgaaaca acgaaagcgac caacttgggt gaaaacaaaa    1560
gctgaagcaa ctacttttac aacaacggct gatggattag ttgatatcac agggcttaaa    1620
tacggtacct attatttaga agaaactgta gctcctgatg attatgtctt gttaacaaat    1680
cggattgaat ttgtggtcaa tgaacaatca tatggcacaa cagaaaacct agtttcacca    1740
gaaaaagtac caaacaaaca caaaggtacc ttaccttcaa caggtggcaa aggaatctac    1800
gtttacttag gaagtggcgc agtcttgcta cttattgcag gagtctactt tgctagacgt    1860
agaaaagaaa atgct                                                     1875
```

Seq ID 29
```
atgaaaacgc caacgtttgg aaaaagagaa gagacattga cctatcaaac aagatatgct     60
gcatacatta ttgtttcaaa gccagaaaat aacacaatgg ttctggtaca agcgccaaat    120
ggagcttatt ttttaccagg aggcgaaatt gaaggcaccg aaacgaaaga agaagcgatc    180
caccgagaag tgctagaaga attaggaatt agtgtagaaa ttggctgtta tttaggggaa    240
gctgatgaat atttttattc gaaccatcgc caaactgcct attataatcc tggatatttt    300
tatgtagcga atacttggcg gcaacttagc gaacccttag aaaggaccaa tactttacac    360
tgggttgcac cagaagaagc agtccgttta ttaaaaagag gaagccaccg ttgggccgtt    420
```

```
gaaaaatggc tagcggcagc atca                                           444
```

Seq ID 30
```
atggcacgcg tagaaagttt tgaattagat cacaacacag taaaagcacc atatgttcgc    60
cttgctggca cagaacaaaa tggtgatgcg ttagtcgaaa aatatgactt acgtttctta   120
caaccaaaca aagatgccct accaacaggc gcattacaca cgttggaaca tttattagca   180
gttaacatgc gtgatgaatt aaaaggaatc attgacattt cgccaatggg ttgccgcact   240
ggttttttata tgattatgtg ggatcaacat tcaccacaag aaatccgtga tgcattagtc   300
aacgttttaa acaaagtaat caatacgaaa gttgttccag cagtctctgc aaaagagtgc   360
ggaaactaca aagatcattc tttatttgca gcgaaagaat acgcaaaaat cgtcttagac   420
caaggaatta gtttagatcc atttgaacgt attctg                             456
```

Seq ID 31
```
atgaacattt ttgctgtccg tttaaaagag gctttaacag ccaagaacat caaacctagt    60
gatttagcga aaaaaactgg catcggtaaa tcttcgatca gcgattggct agctggtcgt   120
tacgaagcaa aacaagacaa agtttatcgc attgcagatg cattagacat taatgaggcc   180
tggttgatgg gacaagaagt tcccatggaa aaaaatgcct caactatcga ccgcatttat   240
aaaaaattag aaccccaacg acaagccatc gtttatcaat ttgccgaaca acaattacac   300
gaacaacaaa cgcaagcaga aattctctca ttccctcgcc gtgacgaaat gacactggct   360
gcccacgctg gggatccaga aaagatattt tcaaaagaag agatcgagaa aattcacgat   420
tacttggatg aaaattgatgc caaatatcaa caatcgattt cttctgacaa aaaagaggat   480
```

Seq ID 32
```
atgtataaat ttgttgatac caatcaagca actcattcaa cgcctcttcc ttcagaagcg    60
ttgaatttta acggccaatt tttagaaaaa gtcatccctg gctatcaaac attatcagtt   120
tcaggacgag aattagttcc aagcgaaatt gaaagctatc aattagggat tcgtgatggt   180
aaacgtcacg tttatgcgcg aattccagaa cgagaattaa cagtcaaata tcgcctttca   240
gctgtgaata atgaagcatt tcgagatgca tttaatcatt taaacgttgc tttgtttacg   300
gaaaaagacg tttctatttg gtttaacgat gaaccggaaa tgctgatggt tggcagtaag   360
tcttcagtga gtgatgtacc cgaaggtgtt aaccaagtaa caggcaccct tactttattg   420
ctttctgatc cgtataaata cacacggagt gatgcgacta gtgtgatgtg gggttcgcca   480
accattacat ttcaagcgaa ttacttaatg gggaatacag gctcaggtgc atttgatttt   540
ccaatttta ttgaaggcgg ggcttattgg ggatcaacca tgattacctt tcaaaatcgg   600
gcttacacga tggggggattt aggcaaagaa gttcggccaa ttgaaatttta tcctacggtt   660
gaaggattaa aagtcaaacc gaccattatt ttaacaggaa ccggacgtgg tgtttggatt   720
aaaacacgga acgatacaat taacttagga gactttgatc gttcggaaat tattatcgat   780
actgaaaatt tttatctgac aaaaaatggt gcaccgatga ttcgaccaat gaacgatttt   840
tatctatatc ccaatgaacc gctgtatatt caagccaaag atagcgactt ccgcttgacg   900
attcgctatc ctaaccgatt tgtg                                          924
```

Seq ID 33
```
atggattttt tcagtagtga agaaatagtt gcaagcttag gtgataaggc agaaatgaaa    60
gggaaattat cgtttgttcg actctgtatt ttaggcatca tggctggttt ttgtattgca   120
ctgggttact tggcgtttat tcgtatcact ggaactgccc ccacagaatg gggaagtttt   180
aacaacttc tgggtggagc tttatttcca gttgggttga tagcgcttac ttttgttggt   240
ggtgagttag caacaggcaa tatgatggta atgacattag gcgtattaca gaaaaaggtt   300
cgagtgggcg ctctttgtta taactggata gttgttttat tgaccaattg tctaggcgga   360
tttcttgtgg cgtatttatt tggccatatt gttgggttaga ctgaaggtgg gtttctggat   420
aaaacagtcg cagttgctca ggcgaaaatt gctgatccac cgctagttgc ttttgtttca   480
gggattggct gtaatatttt tgtttgttta gctgtttatc tcggagcgtt ggcaaaatca   540
tatttaggta aaatgttttg cttatggttt cctgtcatgg tttttcgttgt ttgtgggttt   600
caacacgtgg tcgctaatgc atttattatt ccagcggcaa ttttttcaca aagtaccaca   660
atcagttggt gggattattt gcagaataca ttatgggtat ttttaggcaa tgcagtaggt   720
ggcagtcttt ttatggccgt accattgatc tttatgacaa aaccagcaac tgtcaaacca   780
agagttgaga aaacgattca gacggaggag ctctatggca at                      822
```

Seq ID 34
```
atgaatcagc gtattctatc aacattaggg tttgataaag taaaacaaca gttgttgcaa    60
tttattgtga cagctcaggg gacaaacgaa gtttcggaat tattaccaat tgctgatgaa   120
aataaaattc aatcatggct aaatgaaaca caggatggtt taaaagttca acgattacgt   180
ggtggaatcc ccattcctaa attagaaaat atccaaccac acatgaaacg gattgaaatt   240
ggtgcggatt aaatggcat agaattagca caagtaggcc gtgtcttgtc gactacttca   300
gaactgactc gttttttttga tgagtaagt gaaaatgaag ttgattttga acgtttatat   360
atgtggcgag agcaactaga agttttgcct gaattaaatc gccaactcaa acaagccatt   420
gatgatgatg gctacgtcac agatgaagct tcgccagcct taaaggcaat tcggcaaaat   480
attcgtcgta gtgagcagac cattcgagaa gaactagata gtattattcg tgggaaaaac   540
gcacgttatc tgagtgatgc ccttgttact atgcgaaatg aacgctacgt aattccagtc   600
aaacaagaat ataaaaatat ttttggtggc gtggttcacg accaaagtgc ttcaggacaa   660
accttattta ttgaacctaa acaaatttta gaaatgaata atcgtttgcg acaacaacaa   720
attgccgaaa gaaatgagat tactcgcatc ttggcagaac tgtcggctga gttggttcct   780
tatcgccgag aaattacgca taatgcttat gtcattggta aacttgattt tattaatgct   840
```

```
aaggcgcgac taggaaaaga actaaaagca gttgttccag aaattagtca agcaaatcat    900
gtagtttta aacaagcccg gcacccacta ttaaatccag aaaaagctgt agccaatgac    960
attgtgattg gagaagaata tcaagcaatt gtgattacag ggcctaatac cggcgggaaa   1020
accattactc tgaaaacgtt aggtctttta caattaatgg gtcaagctgg attacctatt   1080
ccagttgaag aagagagtaa aatggggatt tttacagaag tatttgcgga tatcgggat    1140
gaacaatcga ttgaacaaag tttaagtact ttctcttcac atatgacgaa tatcgtgtct   1200
gtcttaaaga aagtcgatca tcaaagttta gtactatttg atgaattagg tgctgggaca   1260
gatccgcaag aaggggccgc tttagcgatt gccattttag attcattggg cgctaaagga   1320
gcctatgtga tggcaacaac ccattatcct gaattaaaag tgtatggcta taatcgagct   1380
ggaacaatca atgccagtat ggaatttgat gtagatactt tgagtccaac ctatcgttta   1440
ttaattggcg tgcctggccg aagtaatgct tttgaaattt cgaaacgtct tggattagac   1500
aacagtatta ttgaggctgc aaaacaaata atggacggtg aaagtcaaga tttaaacgaa   1560
atgattgaag acttggaaaa ccgtcgcaaa atggccgaaa ctgaatattt agaagctcgc   1620
cactatgtcg atgaatcggc cgctttgcat aaggaactga aagaagccta tcaagtattt   1680
tttgaagaac gagaaaaaga attacaaaaa gcacggaaag aagcaaataa aattattgca   1740
gaagcggaag aaaacgcaga aaccattatt tctgatattc gcaaaatgca attggaaagt   1800
ggccaacaag gtggtgtcaa agaacatcag ttaattgatg cgaaaacaca actttctcaa   1860
ttacaccatg aagaaaccaa gcttgctaaa aacaaagtgc tgaagaaggc gaaagaacag   1920
aaaaaattaa aagctggcga tgaagtaatt gtcaatactt acggccaacg agggacctta   1980
ttaaaagaca acggcaaagg acaatggcaa gtccaattgg gcattttaaa aatgaacgta   2040
tctgaagagg atatgacgcc tgtcgctcca caaaaagaag cgaaaccaag agttacaacg   2100
gttcgttctg cagaaagcag tcatgttggc acacaactag acttacgtgg caaacgctac   2160
gaagaagcat tggcagaagt tgaccaatac attgatgctg cgattctagc gggctatcca   2220
caagtaacca ttgttcacgt caaaggaaca ggtgcattaa gaaccggaat aacggaattt   2280
ttaaaaaatc accgcagtgt gaaaagctat gaatttgcgc cacaaaacca aggtggtaat   2340
ggtgcgactg tcgtaaaatt ccaa                                           2364
```

Seq ID 35

```
atgaaagaaa taactggagc cactcgttta gctgggctat tcgcgaaacc cagccaacac     60
agtatttcac cgttgattca taatacagca tttcaaaatt taggagttga tgctcggtat    120
ctggcgtttg acgttggaca agagacattg ccacaagcaa ttgaagcgat tcgaacgttt    180
cacatgttag gggccaactt atcaatgccc aataaagtgg cggctgtaag ttatatggat    240
gaactaagtc ctaccgctca actggttggc gcaattaata cgattgtcaa caaagatgga    300
aaactttacg gagacagcac ggatggtact ggtttttatgt ggagtttgaa agagaaaaag    360
gttgacgttt tcagaataa aatgaccatc ttaggaacag gtggtgcagc cttatcaatc    420
attgcccaag ctgcttaga tggcgtgaaa gaaatcgccg tttacaacag gaaaagccg    480
ggctttaacg acagtcaaaa aaaactggca aatttcactg aacgaaccaa ctgtgtaatt    540
catttaaacg atttagcgga tactgaaaaa ctagcaaaag atgttgctga aagcgtcttg    600
ttagttaatg caacgagtgt gggtatgcat ccacatgcgc atagtagtcc tatagaaaat    660
tatgcaatga ttcaaccgaa gttatttgtg tatgatgcta tttataatcc cagagaaaca    720
cagttattaa aagaagcccg tttacgtggt gcagaaacaa gcaacggctt ggacatgcta    780
ctttatcaag gcgctgctgc ttttgaacaa tggacaggac aaaaaatgcc tgtatcagtc    840
gtaaaacgta aaattgaaaa taga                                           864
```

Seq ID 36

```
atgactaaag gttatgtaaa atcagtaaca gagttaattg gtcaaagccc agttgtgaaa     60
ttaaaacgga tggtaccaga aggagctgca gatgtttttg ttaaactaga gttctttaat    120
cctggtggca gtgtgaaaga tcgaattgct ttaagcatga tccaacaagc cgaagcagat    180
ggtcgcctaa aacctgaca aacgattatt gagccgacta gtggtaatac aggcattggg    240
ttagcaatgg taggtgctgc gttggggatt ccagttaaga ttgtgatgcc agatactatg    300
agtattgaac gccgcaaatt aatgcaagca tatggggctg atctattatt aacccctggt    360
gccgaaggaa tgaaagggc aattgcaaaa gctacagcat tagcagaaga acacgggtac    420
ttcatgcctt tgcaatttaa taatccagct aatccaatgg tacatgaaca aaaaacagga    480
aaagaaattg ttgatgtctt tggtaaacgt ggcttagatg cgtttgtttc tggtgttggc    540
actggaggaa ccgttacagg agttggccat gaattgaaac ggattttcc agatactatg    600
attgttgcag tagaaccaac agagtcgcct gttttagaag gtggcgaacc aggtccacat    660
aaaatccaag gaataggcgc tggttttgtc ccagaagttt tagacaccac cgtttatcaa    720
aaagttgccg ctgttctag tgaagacgca ttagaaactg ctcgtttaat ggggccaaaa    780
gaaggtatcc ttgtggggat ttcagcaggg gcagcaatta aagctgccat tgatttggca    840
gtcgaattag gcgcaggcaa acgtgtctta gcgctggttc cggataacgg tgaacgttat    900
ctttcgacag ctctttacga atttccagaa                                     930
```

Seq ID 37

```
gtgaaaaatc aacatttaac tacgtcacaa ggaagtccag ttggcgataa tcaaaattcg     60
ctaactgcag gagaatttgg tcctgtccta atccaagacg ttcatttatt agaaaaatta    120
gctcatttta atcgtgagcg cgtaccagaa cgggttgttc atgccaaagg tgctggtgct    180
catgggattt tcaaggtgag ccaatcaatg gcacaatata ctaaagccga ttttttatct    240
gaagttggca aagaaacacc tttattcgct cgttttcaa ctgtagctgg agagttaggt    300
tcatcagata cattgcgtga tcctcgcggt tttgcgctta aattctatac agacgaagga    360
aattacgact tagttggcaa taatacgccc atttttttca ttcgtgatgc aattaagttt    420
ccagatttta ttcatagcca aaaaagaaat ccgcggacgc atttgaaaag tccagaagct    480
```

```
gtttgggatt tttggtctca ttcaccggaa agtcttcatc aagtaacaat tttaatgagt    540
gatcgtggca ttccgttgtc gttccgtcat atgcatggtt ttggtagcca tacattcaaa    600
tgggtcaatg ctgctggaga agtatttttt gttaaatatc acttcaagac gaatcaagga    660
atcaaaaatc tagagagtca attagcagaa gaaattgctg ggaaaaaccc agatttccat    720
attgaagatt tgcataatgc aattgaaaat caagaatttc cttcttggac attatctgtg    780
caaattatcc cgtatgcaga tgcgttaaca atgaaagaaa cacttttgta tgtaacaaaa    840
acggtttctc aaaaagagta tccgctgatt gaagttggca cgatgacttt aaatagaaat    900
ccagagaatt attttgcaga ggttgaacaa gtaaccttt caccaggaa tttcgttcct    960
ggtattgaag cttctccgga taattattg caaggacgtt tgttcgctta tggggacgca   1020
catcgtcatc gcgtggccgc taatagtcac caattgccaa tcaaccaagc aaaagcgcct   1080
gtaaataact accaaaaaga cggcaatatg cgttttaaca atggcaatag cgaaattaat   1140
tatgaaccaa atagttatac tgaaacacca aaagaagatc ctacagcgaa aattagttcc   1200
tttgaagttg aaggaaatgt tggtaattat agctataatc aagatcactt tacacaagca   1260
aacgctttgt ataatttgct gccaagcgaa gaaaaagaaa acttaattaa caatatagcg   1320
gcatctttag gtcaagtgaa aaatcaagaa atcatcgcgc gtcaaattga tttgttcact   1380
agagtaaatc cagaatatgg agcacgtgtc gcacaagcca tcaagcaaca agca         1434
```

Seq ID 38
```
atggcagaaa attatcagca ggcagccaaa gacattattc aattaatcgg tatggataac     60
attatatcgg tcacccattg tcaaacgcgg ttacgtttta ttttgaagga ccatgaacaa    120
gtagatggta aacaactaga aaaaatcgat ttagtcaagg gggttttta taatggtggt    180
caataccaag tgattttagg aactggtatt gtgactaaag tgtatgatga aatagaaaaa    240
cttggaatca atgttgtgtc aaaagcagag caaacagcaa ttttaaaaaa taatgaaaca    300
ggcatgcgca aaacaatgcg cattttaagt gaaatattca ttccaattgt tccagtaatt    360
gcagcaacag gattgttcct agggttaaaa ggtgttattt ttaacgatac tttttacaa    420
ttattcggag caagtgtagc caatattcct gaaagctttc aacaaattgt ctctgtcatt    480
acagataccg tctttgcctt tttgccagct ttgattgttt ggtctacctt taaagcattt    540
aatgcgacgc ctgtaatcgc gattgttatt ggtttaatga tggtatcacc catttttacca    600
aacgcttatg cagttgcaac accagattct ggcgtgaaag ctattatggc ttttggtttt    660
atcccagttg tcggtgcaca agggtccgtt ttaagtgcaa ttgctgctgg tataattggc    720
gcaaagattg aacttttttt ccgaaaaaag atgccgaata ttcttgatca aattttttaca    780
ccctttatga ctatgttaat tacttttta atcatgattc tcggaattgg cccaatttttg    840
cataccgttg agttaggcat ggtagatgtt gtccaatggt tgattgggtt gccctttagga    900
ttaggtggtt tcgtgattgg tgcctcatat cctttaatgg ttttgattgg tattcatcat    960
acactaacta tggtagaaac atcattatta gcaaatacag gatttaatgc attgattaca   1020
atatgtgcca tgtatggggt tgccaacgtt ggcagttgtt tagcctttgc gaaaaaagcg   1080
caagatagca aagtgaagtc gacagccatt ggttctatgt tgtctcaatt atttggagtt   1140
agtgagcctg tgttatttgg attacttatt cgttggaact taaaacgct gctatgtgtc   1200
ttgtttacat caggtttagg cggagcaatt ttagctattt ttcatattca atcaaattct   1260
tatggcttag ccgttattcc ttcttttctta atgtatatct atagtgcaca tcaattagtg   1320
atctatttat tagtcgcact tttatctgtc ggtgtatgtt atgcattaac cagtctattt   1380
gcaattccgc aagaagtttt gatttcagat aaagtaatcg aagaggaaga acgcgaagtc   1440
tttgaaatgc aacacaatac cttagatgaa caacttttttt ctccggtaac tggttacgcc   1500
atgaatctta ctgctgttaa tgacccagta ttttcaagcg agatgatggg caaaggttta   1560
gcaatcatgc caacagcgaa caaggtttac gcgcctgcag atggtttatt aaacttggtt   1620
gcagaaactg ggcacgctta tggtatccaa acggacgcag gagctgaagt attaattcat   1680
ataggaattg atactgtcac atgggacag gaggtctttc agacacaagt aacccaaggt   1740
catcgtgtta aaaaggaga tcttttagga acgttcgatc gaaaagccat taaagaagct   1800
gggcttgact caactgtcat ggttatcatt acaaatactt caagttattt aagcgtagag   1860
ccaatgatgt cagatcataa tgaaattacg ccggaacaaa tcatttttaaa tctaaataca   1920
cctaac                                                             1926
```

Seq ID 39
```
gtggaattag tagataaaga cttacgttcc atccaagaaa caagaaattt aatcagaaaa     60
gccaaagagg cacagcaagt attagcaaca ttttcgcaaa aacaaatcga tgcgattgtt    120
caagcggttt cagaggctac atttaaccaa cgagaaaagt tagcgaaaat ggccaatgaa    180
gaaacaggat ttggtatcta tgaagataaa atcatcaaaa atgctttcgc ttcaaaagta    240
gtatatgacg aaatgaagga caaagctaca gttggtgtga tccatgatga tgctgccaaa    300
aaagtcactg aaattgctgt cccagtgggt gtaattgcag gattaattcc ttctaccaac    360
ccgacatcaa cagtcattta caaagcatta atttcattaa aagctgccaa cagcattgtt    420
ttttcaccac atcccaatgc gttgaaatcg attattgaaa cagtgaaaat tatccaaaaa    480
gcggctattg ctgcaggggc accagaaggt tgtgtctcgg tcatcaaaac gccaaccatg    540
caagcgacca gcgaattaat gaaaaataaa gaaacaaatt taattttagc tactggtgga    600
aacgcaatgg ttaaagccgc gtactcttct ggaacaccag caatcggtgt tgggcctggt    660
aatggaccag cctacatcga acgaagcgca aatgttcctc atgcagtgaa acaaataatg    720
gattcaaaaa ctttcgataa cggtacaatt tgtgcatcgg aacaatcaat tattgttgaa    780
actgtcaacc gtgaagcggt caaagaagag ttaattaagc agggagcgta tttcttaagt    840
ccagccgaag ccgacaaatt agctaaattt attttacgac caaacggtac aatgaacccg    900
caaatcgtgg gacgttccgt tcaacatatt gcttcattag ttggtttatc tattccaaaa    960
gatcgccgct taattgttgc cgaagaaaca catgtaggct tgaaatatcc attttcaaga   1020
gaaaaattgg caccgattat tgctttcctat acagttgaaa actgggaagc agcttgcgcg   1080
```

```
ttatcaattg aaattttgaa aggtgaaggt gcgggtcata caatgggcat tcacacagaa   1140
aacaaagaag tcattcgtga atttggctta agaaaacccg tctctcgttt gcttgttaac   1200
acctctggta cacttggcgg cattggtgct tcaactaact tagtgccagc attaacactt   1260
ggctgtggcg cagttggtgg cagctcaact tctgataata ttggggttga aaatctgttc   1320
aatttacgtc gtgttgcata tggcgtgcgt gatttagaag agattcgtca agaatttggt   1380
caaacatcaa ccacatctgt ggctacttct tgcgaaacaa ccaatcaaga agaactagta   1440
aatgccgtgg tagctcaagt tttagctcgc ttaaac                           1476


Seq ID 40
atgaatgaat taaacaaaac accaaaagaa atagtcaaag aactagatca atatattgtt    60
ggacagcaag ctgccaaaaa atcagtggcg gtagctttac gtaaccgcta tcgtcgcttg   120
caattagaag aaaaatatgca acaagatatt acgcctaaaa acttactaat gattggacca   180
acaggtgttg gtaaaactga gattgctcgt cgtctcgcaa aaattgtcaa tgcgcctttt   240
gtaaaagttg aagcaaccaa atttacagaa gtaggctatg ttggtcggga tgtcgaatca   300
atggtgcggg atttagtaga aaatgcgatc caaattgttg aaaaacaaca atacagtcgc   360
gtgtatgcgc aagcattaaa aaaggccaat caacgtttag tcaaagtatt agtgcctgga   420
attaaaaaag aacaaaaaca ggctggcggt aatcaatttg aacaaatgat gaacatgttt   480
aacatggctc agcaacagca agaagcacaa gaagaagtaa cggaagatat tcgaacgaat   540
cgccgaacaa ttttggaaca gctggaaaaa ggtctattag ataatcgtga agtaacaatt   600
gaaattgaag aaccgaagaa aacgatgcca gctatgaaca atggcctaga acaaatgggt   660
atcgacttaa atgaaacgtt aggcgctctg tcaccaaaga aaaaaatcga acgtactgta   720
acggtgaaag aagcacaaga attattagtg aaagaagaat cagcaaaaat tgtcaatgac   780
gctgatattc atagtgaagc tattcgttta gctgaatcaa gcggaattat ttttatcgat   840
gagattgata aaatcacctc taaaagtcaa caaaattcgg gcgaagtctc tcgtgaagga   900
gtacaaagag atattttgcc gattgttgaa ggctcccaag ttaacacgaa gtatggtcac   960
ttacaaacgg atcacatttt atttatcgct tcaggtgctt tccacttgtc aaaaccaagt   1020
gacttgattc cagaattaca aggccgcttc ccaattcgag ttgaattaga tgatttaacg   1080
gcggatgact tcgtaagtat cttaactgag ccaaacaatg ctttaattaa acaatatgta   1140
gcattaattg gcacagaaaa tgtttcagtc atctttacaa aagaagcaat tgaacggtta   1200
gcacacatcg cttatgatgt aaaccgtgat acagataata ttggggcgcg tcgtttacac   1260
acaatttttag agcgtttatt agaagatttta ttatatgaag caccagatat gcaaatgggt   1320
gaaattacga ttaccgaagc atatgtcaat gaaaaattga atgacattgt tcaaaatgaa   1380
gatttaagtc gttacatctt a                                           1401


Seq ID 41
atggcaaaac gaattatcat tatgaacttt gatattgaaa gtaaatcgta ccaagcattt    60
tccgaaataa aaaagatgca ggcagaaaga cagcttaaag gtgagcaaat ggcggttgtt   120
acgcatgtta atgatggcca gcatcaattt aaaatcaatg attttatcga ttttacaggc   180
aacaatcata catcaaaaga tagcatgatt gggatgctag taggcatatt aggtggtcct   240
ctgggcattc tttttggctg gtttgctgga agtatgtatg gtgcaagcaa agacgccaaa   300
gaaattcaag aagcacaaac ggttttttgaa catgtgattc aaaagattga tgaaggacaa   360
acgggattgt tattaattgc agaagaagaa gacaaccgtc cgctcaacca attggttatg   420
tttgacttag gtggcgaaat cacgcggctt gatttagagg aagtccaaca agaaattaac   480
gatgcgaacg aagttgcaaa tgaagcgaaa caatcgttgg caagcaaaaa agaacaacac   540
aaagaagcaa catcgaaaga agaa                                         564


Seq ID 42
atggaaaaac aaacaattac aatttatgat gttgctagag aggcaaatgt atctatggct    60
actgtttctc gtgttgtcaa tggtaatccc aatgtaaaac cagcaacgcg taaaaaagtc   120
ttagaagtga ttgatcgctt agattaccgt cctaacgcag tcgcacgtgg tttagcaagt   180
aaaaaaacga caacagtggg tgtcattatt ccagatgtta gtaatgcatt ttttgcttca   240
ttagcacgtg gcattgacga cgtagctaca atgtataaat acaatatcat tttagctaat   300
tcagatggag acgatcaaaa agaagttact gtattaaaca atttacttgc caagcaagta   360
gatgggatta tctttatggg tcatcgcatt acagatgata ttcgtggtga atttttcacgt   420
tcaaaaaacac ctgttgtctt agcaggttca attgatccag acgaacaagt tggtagcgtt   480
aatattgatt acacagaagc aacgaaagat gcaacagcga ctttagcgaa aaatgggaac   540
aaaaaaattg cttttgttag tggtgcatta atcgatccaa ttaatggtca aaaccgaatg   600
aaaggctaca aagaagcttt agccgaaaat ggcttatctt ataatgaagg cttagttttt   660
gaatctgaat ataaatttaa agcaggcatt aatttggctg agcgtgtccg taatagtggg   720
gcaactgctg cctttgttac agatgatgag ctagcaattg gcttattaga tgggatgcta   780
gatgctggtg tgaaagttcc agaagaattt gaaatcatta caagcaataa ctcattgctt   840
acagaagtct ctcgtccacg tctatcaagt attacacaac ctttatatga tatcggtgcc   900
gtgtcaatgc gtttattaac aaaactaatg aacaaagaag aaattgaaga aaaaacagtt   960
gtttttacctt atggaattga tcaaaaaggt tcaacaaaa                        999


Seq ID 43
ttgaagaaat acaatgttga ttggaattac tgggttgtcc gcttttttatt tgtaatggcg    60
ctgattgttg gctatttatt aatcacgaat tatcagcact ttgtccatag tgtttcgggt   120
ttattaggca tcttatcccc attcattact gggtttgtga ttgcttattt attgagtgga   180
agtcagaaaa aaatcgaggg attactagaa agagttcctt taccagtcgt taaaaaagcg   240
aaacacggct tgagcgtact ccttttgtac ctgattattt tgttattttt tgttttgaca   300
```

```
ttaaactata tcgttccgct actcattagt aatttagtag atttagcaaa ctctttaccc      360
actttttatg atcacatggt tcagtttgta atgagtttag aagataaagg gattttaaaa      420
acagccgcga ttgaaaaata tttaaatagt gtccttaaag atttgtcgcc agaacgtttt      480
ttaaatcaat ggacacaagc gttgttttca ttaggaacgt tgactaaaaa tgtttcatca      540
ttcttttaa atgcattttt gacgttgatt atctcaatat atgcgctggt atttaagcaa      600
tctattttga catttgttga aaaggctgcc cacaaattgc tgtctgaaaa agtgtacaag      660
caaacacaaa cttggttaaa tacaacgaat aaaatttct ataaatttat cagctgccaa      720
tttcttgatg cttgtataat tggcgtttca tcaacgattt tgttaagtat tttgaatgtg      780
aaatttgctg tgactttggg gattttatta ggtatttgta atatgattcc ttattttggt      840
tcgattttg cgtcaattgt tgctggtgtg attacgcttt tcacaggtgg tgttactcaa      900
gcaattacgg tgttactagt tttattaatt ttacaacgaa ttgacggcaa tattattggt      960
cctcgaatta tgggagatgc cctaaatgtc aatccgattt taattattgt ttcaatcacg     1020
attgggggag cgtacttcgg cgtcttaggg atgtttttag cagttcctgt tgctgcaatc     1080
attaaaaatta ttgtctcgga gtggctaaat gaatcaaaag aaaatgataa gattgtagat     1140
tctattgaat cc                                                         1152
```

Seq ID 44

```
atgaaaaagt tctccatacg aaaaattagt gctggttttt tgtttctgat tttagtaact       60
ttgatcgccg gttttagctt gtctgcaaat gcagaagagt atatcgttcc tgccgaaagt      120
cattcacgac aaaaaagatc gttactggac cctgaggaca gaagacaaga agtggcagat      180
acaaccgaag cgccttttgc gtcaatcgga agaatcattc ccctgccag taaaccaggc      240
tatatttctt taggaacagg ctttgttgtt ggaaccaata caattgtcac caataatcat      300
gtggctgaaa gttttaagaa tgccaaagta ttaaatccga atgccaaaga tgatgcttgg      360
ttttatccag gtcgagatgg cagtgcgaca ccattggca aattcaaagt gattgatgta      420
gctttttccc cgaatgcgga tattgcggta gtgactgtcg gcaaacaaaa cgatcgtcca      480
gatggcccag agttgggaga aattttaacg ccatttgttt tgaaaaagtt tgaatcttca      540
gatacccatg tcacaatatc aggctatcca ggtgagaaaa accacacaca atggtctcat      600
gaaaatgatt tgtttacatc taactttaca gacttagaaa atccattact attttatgat      660
atcgatacaa ccggcggtca atctggttca ccaatctata atgatcaggt tgaagtagtt      720
ggtgttcatt ccaatggcgg cattaagcaa acaggaaatc atggtcaaag actaaatgaa      780
gtgaattata actttattgt taatcgagtg aatgaagaag aaaataaacg tttatccgct      840
gtgccagcag cg                                                         852
```

Seq ID 45

```
atgaaaagaa agaaaatcaa gaatcaacta cttgttttaa caagtattct cgttcttctg       60
agtttggaag tcgctccagt tgttacgttg gctgcagaat taccaagtag ctcgctacaa      120
acagctctgt cctcggaaac gacaattact tcggaagaaa aagtaaccga aacaaccgaa      180
acaactgttg ctacatcgac aacgtccact agtagttcgt catctgaatc aagtagttct      240
accgatacga ctaccgaatc aacgagtcag tcaacgacag aaacgacaac cacaaatact      300
agttctgaaa caaaaaaaga acctacagag ccaacggtat catcagaaat cacgcagcct      360
gttgaacaat cgcaaccacc acaagttcct gtgaccaagc aagaacccga agaacctatt      420
caagttcctg aagccaacaa taactttgtt gaagaaaatc aggctgtatc gttaaatccg      480
tccctcaaag tagatgaaat agcttcttca aatttaaaag gctacgaact tccactgttg      540
tcgtcttttg gcgaaaagaa acgggcagtc gtggtcgctg aagcattaag cacgttgggg      600
aaaacaaaaa aagaatttaa tttgacggag caagcattga ccagtagttt ccttgcccaa      660
aaaatttacc aacaattatt taaaatagac attggcagta cgccgcaaga acaaatgaca      720
tttggtaaag ttcgttctat agaggaggca gaacctggtg atctgatttt ttggcaaaca      780
gctgaaggaa aaacgcttca aaatggtgtt tatcttggcc aaggaaaata tttaattgcc      840
gcagcggaat cggattctaa agagaaacca gaagtaatcg ctcagttaga gaatatctat      900
actgccaaac aacaaccaga ctcaaaagag gaaaaacgct tggtagtgac taatcctttt      960
aaggagttca tattaacgga gtatggcaaa gaagttttag caacatatgg tgcgtcattt     1020
gaaatgcaga agtcagaaca aaccacagcg tttattaaaa aaattggtga gacagccaga     1080
gaacttggcg aaaaatacga tgtttttgca tcagtaatga tagctcaggc tatttttagaa     1140
agtggctctg gcgaaagtca actagccaaa gaaccttatt ataatctctt tggcgtgaaa     1200
ggttcgttcc aaggaaatag tgttagcttt tcaacaaaag aggctgatca gaggggccaa     1260
ctttatacaa tttccgctgg gtttcgcgat tacggtggct ataatgactc actgcaagat     1320
tatgttcaat tactccgaca agggattgac ggaaatcaag atttttataa accagcgtgg     1380
cgttcggaag ccaaaaacta ccttcaagca acgcgttttt taacaggtaa atatgcaaca     1440
gacaagcaat acgataataa attaaaattcg ttaattgcgg tttataactt aacccaattt     1500
gatttaccaa aaacagtgga cggtttgatt atccaatcta aaaataagct gtcggaagcg     1560
gaacagcaac aaatgcattt tccagtctat gatgggatca attatatcg aagtggtagc     1620
tatcccgtgg gccaatgtac gtggtatgtc tataatcgtt tcaaacagtt gggcacttct     1680
gtggatgaat ttatgggaaa tggcagtgac tggggaagaa aaggacgtgc attaggctat     1740
caagttagtt cattgcctaa agctggtcgt gcgattagtt tccaaccagg agttgctggc     1800
gcagataatc aatatggaca tgttgctttt gttgaagcgg tcacttctga tggaatcatt     1860
atttctgaaa gcaatgtcat caatgatcaa acgatttcgt atcgtgtctt acctaatgta     1920
attgcttata gtagtggcgt gacctacatt ggcgct                               1956
```

Seq ID 46

```
gtggatgcag aagcttattt aggtcctgat gatgaggcgg cccatagctt tggcttgaga       60
aaaacgccac gcttacaagc catgtatgac aaaccaggta cgatttattt atatactatg      120
```

```
cacacgcatt taattttaaa tatggtgaca caagaacaag gcaaaccaca aggggtgatg    180
atccgcgcga ttgaacctgt tgaaggtgtc gacaaaatga ttgaaaatcg ccaaggacgt    240
caaggcgtgg aattgaccaa tggaccaggt aaattagttg cagccttagg aattgataag    300
caattatatg ggcagtcgat tttttctagt tcgttgccag tcgtgccaga aaaacgaaaa    360
tttccgaaaa aaattgaggc actaccacga attgggattc ccaataaagg tcgttcggaca    420
gagttgcctt taagatacgt cgtggctggc aatccatata tttctaaaca aaaaagaaca    480
gcagtagacc aaatagattt tggctggaag gatgaagaaa atgaaaaaag caacaatgct    540
cacatactta gaggaacaac t                                             561


Seq ID 47
atggcaaaaa agaaaaatgt aagtgtcatt tctgtagaga agccaacgtg gttcccacta     60
acagacgaaa cgggcgcttt tccagtttac ggagcgccaa ttacaatcgg tactgctgta    120
agtatcaaac cagatgttac aacagaaaca acgcctgact atggcgatag tgtagttcaa    180
gatcagtatg ttgcatttgg tggtgcagaa gttactttag aaacaaacgg ctaccaaaat    240
gaagtttttg ctgaaattac agggggaaaa aaattaaagg gtggcgtatt gcggtctgcg    300
gatgatattg catcagatgg cgcatttgct tatcgtcgcc gaaaatcgaa cggtaaatat    360
cgctatacga ttttctataa aggaaaattt gcttaaacat ctgatgaaac atctacatta    420
gaaggaagtt cagtatctta tacccatcca gaatggacgg ggtctttcgt tgatgttcca    480
gggttggggtt atatgtattc cgtggatgaa gacgatgaag gtgttgactt agagatgatt    540
aaaaactggt ttactgaggt aatggatcca cgtaaagaaa atactactgc tgttactggt    600
gtaacttttag accaaacaga gttaaattta aaagttggcc aaacagcaac cttaacaccg    660
acaattacac cagataaacgc ctcaaataaa aaatatcagt tccgttcaga aagtgaggct    720
attggaactg taacaccaat tcaagggaag gttactgctg taggagaagg gacaacggaa    780
atcgtagtca caacagaaga tggtaacttt accgcaaaat gtacattaaa tgtaacaaca    840
gcagat                                                              846


Seq ID 48
atgtttcttg gcctgatttt tattttattc ggtgtctttg gactctttcg cttagggttt     60
ttaggcacgt tattagcgaa ttgttacgc ttagttgtag gtaataccct tccatttgca    120
gcgatttttac taatcctgta tggcttatta gtcatgattt atgggaaaga ttttccgtta    180
aaaagagggc gccctatttt tggggctgtt ctgatttata ttagtgtttt gttattttttc    240
catgcctttta tgtttcgaaa tgttagtggc agtcaaccag atattttagg taacacttgg    300
gaatttttgc aaagcgatct aaaagctaac caagtgacgc aaaacgtagg gggcggtatg    360
attggcgcac tgctttatca aggaacgtac ttttttggtgg cacaatttgg tagctattta    420
attgccacct tgctactgtt agcgggtatt ttcttgatga gtatgtggga ttttcagcaa    480
attgttgatc atttccaaag tattcaagat cgcctaagtc atgtttcggc aaaatcacag    540
gcgcgacaag aagaaaaaga ggcgaagcga gcagcgaaga aagaagccaa agctgcagaa    600
cgccaagcaa aaattgaagc cgctgcccaa caaaaattac aagagcggga acgcatggaa    660
caagcggcag ccgagcgctt aacaaaaaca cctgttgaaa cccaccaacc aatggtggaa    720
gaaccggccg caccaacgcc tgttcaaatt gattcgtttc aacagcaaaa tcaggcgatg    780
cctgtaccac caattgcagc aacaaaaccg caacgagagc aggaagaaga ggcggcagat    840
gaagctggtg ttctcgaatt tgaaatttct gaagaagcgg aagatcgtga ttatcaacta    900
ccaccaactg atttgttaga tacaattcaa gcaacggatc aaagtggtga atatgaaaaa    960
atcgaaaaaa atattggcgt cttagaacaa accttcaaga gtttttggtgt tgatgctaaa   1020
gtagtcaaag ctagtttagg accatctgtc acaaaatttg aagttcaacc cgctgttggg   1080
gtaaaagtta gcaaaattgt taacttgacg gatgacattg ctttagcgct tgcggccaaa   1140
gacgttcgta tggaggctcc gattcctggg aaatcactaa tcgggattga agttcctaat   1200
agtgcaatta gtacagtatc cttagagat attgttgacg cacaaccaat tcatccaaaa   1260
aaactattag aagtacctt aggtcgagat atttcagggg tggttcaaac tgcagatcta   1320
tcaaaaatgc cgcatttatt aattgcggga tcgacaggga gtgggaaatc ggttgctatc   1380
aatggcatca tcacaggtat tttgatgcaa gcgaaacctc atgaagtaaa attaatgatg   1440
attgacccga aaatggtaga attgaatgtg tataacggga ttcctcattt attaacgcca   1500
gtggttacta atccgagaaa agcggcccaa gccctgcaaa aagtagttca agaaatggaa   1560
tttcgttatg aaaaatttgc agcaacaggc gtccgaaata tcactggtta caatccaacta   1620
attcaacaga aaaatgcgga agatggcgaa aatcgtccaa ttctaccatt tattgttgtc   1680
attgttgatg agttagctga tttaatgatg gtggcaagta acgaagtaga agatgcgatt   1740
atccgtttag cccaaatggc acgggcagca ggtattcata tgatttttagc gacacaacgg   1800
ccaagtgtgg atgtcattac aggaatcatc aaagcaaatg tcccttctcg aatggctttt   1860
gcggtctcaa gtgggacaga ctcgcgtacc attattgata ccaacggggc ggagaaacta   1920
ttaggtcgag gcgatatgct cttcttacca atgggagaaa ataaaccaat tcgtatccaa   1980
ggggcctta tctctgatca agaagtcgaa cgagtggttg cgtttgtaac ggatcaacaa   2040
gaagctgaat atcaagaaag catgatgcca acagatgaac cgaccacttc aggcggtggc   2100
gaagctcctc aagatgagtt gtttgaagaa gcgaaaaact tggtggttga aatgcaaaca   2160
gctagtattt cattgctaca aagacgtttt agaattggct ataatcgtgc ggctcgctta   2220
gtagatgaat tggaagcaca cggagttatt ggaccatcag aaggtagcaa accaagaaaa   2280
gtctttctac aagcggaatc agaagaggca gcaaccgaga cgccagaaca a            2331


Seq ID 49
atgagaaaaa gctttaactt agctgttcaa gcgttaaccg ttcaatatca aggacggacc     60
gctttaaata atatccacgt tactattccc tccggtaaaa ttactggaat tatcggacca    120
aatggtgctg gaaagtcaac atttattaaa gggttattag gcttgattaa aacaaaggaa    180
```

```
cgtgatgttt tgttaaataa tcaagcgatt gaccaacaaa aaacaaccat tgcctatgta    240
gaacaacgta gcgccttgga tcttagtttc ccaattagcg tttttgaaac ggtcttgcta    300
ggaacctatc caaacttggg actattaaaa cgcccaggaa agaaagaaaa gcaagcagcc    360
atggctggcc taaaaatggt gcaattagaa gactatgcgc aacgccagat tggcgaactt    420
tctggtggcc aattacaacg tgtgtttatc gcccgtgttt tggcccaagg tgctgaggtg    480
attttttttag atgaacccttt cgtcggcatt gatatgtcta gtgaaaaagt gattatggat    540
attcttaaat cattaaaaaa tcaaggtaaa atgattatca ttgttcacca tgatttgcac    600
aaagtgtccc actattttga tgaattaatc gttttgaaaa accggctaat tgctgcgggt    660
cctgttgaac aaacatttac tgcagaaacg ctccaagaag catacggtga tttgttaggt    720
gatttattaa tacagggggt tgcaaaa                                        747


Seq ID 50
atgattggat tagaattgac tttttacagtc cctcaagcag atgaagtcat tcggtcatta    60
gttgaaacat atcatgacga gaaagatatc gtgattggcg cgggcacggt gttagatgct    120
gtgacggctc gcctagctat aatggccgga gccgaattta ttgtaagccc tagctttaat    180
cgggagacag ctgaaatttg taacttgtat caagttcctt acttgccagg atgtatgaca    240
attactgaga tgcagacagc tttgaagagc ggagtagata tcgtgaagtt atttccggca    300
agtacctatg gtccgagtat tatctctgca ttccaagcac ctttgccgta cttgaatatt    360
atgcccacag cgggtgttag tctggataat atggcgtcat ggtttgaggc gggtgtgacg    420
gcagtgggtg ttggcggaaa cttattagct cacgctaaaa ctggtgagtt tggtaaagta    480
acggctttgg caaaacaata tatggccaaa tttagagaaa taaaggaaaa c             531


Seq ID 51
ttggcaggcg taaccaaaga gcagattgca aaagcaaagg aatgggatct gctctcttat    60
ttgcaggcct atgagccgca ggaactgaaa aaaagcggcc cccgcgaata ctgcaccaga    120
acccatgaca gcttaaaaat ctccaacggg aaatggtgct ggaacagccg gggcatcggt    180
gggcgcacgg cccttgatta tctgatcaag gtgcgcggca tggactttgt aggcgcggtg    240
gaaaccctgt gcggctactg tgcgccaccc ccgaaaaag cgccgcccat aaaagaaaaa    300
cagccaaagc ccttccgact gcccccaggcc agccgatgtg catccgctgt tgtgggctat    360
ttgcaggatc gcggcattga cccggaactg cttgcgtgt gcatggaagc cgggattta    420
tacgagagct gcaaatacca gaactgcgtc tttgttgggc gcgacacagc agggaacgcc    480
cgcgccgcca gcctgcgggg gacgcgggac ggcttccgca tggatgtgga gggaagcgac    540
aagcggtaca gcttctgtct gccggcaggc aaggcagact gcccccggct ggcggtggcg    600
gaaagcccta ttgacgcgct ctccctcgcc acattagtaa agctgtccgg cggcgaacca    660
cgggacagcc attatttatc cctgggcgga accggccccc gcgccctgat acagtttctt    720
cacgaccacc cccacgttac ccaagtcagt ctgtgcctcg acaatgacaa ggccggcttg    780
gagggcatgg aacggctcac ccaggaaatc cagaacgatg cggaactttc caggcgggtg    840
aagctggttt accccaaccc gccgaagcag ggcaaggact ataacgagtt tctctcgcacc    900
cacgttaaag cggtacggac ggcacagcgg cagcgggacg cggcgcgg               948


Seq ID 52
gtgaggataa aggagggcg cacggtgggg ctgttcacca aagtttcccc ggaagaaaag    60
gcggcgattg accgcaaaat ggaattgctg ggaacctcca acctgcgggc ttacctacgc    120
aagatggcgg tggacggcta catcgtgcag cttgacatgg cgcggtgct ggaactggta    180
aaattgctgc gttccgtttc cagcaacgtc aaccagatcg cccgccggtg caacagcacc    240
cacaacctct atgagcagga tgtggaggac ttgcggcagg gctactctga aatctggcag    300
ggcgtgaacg acctgctgaa gaaatttgaa gcactg                              336


Seq ID 53
atggcgaata agctacaaac cgtgtcgcag cttgcagacc agaccgccaa ggccgtcacg    60
cggaacgcgg agggctggaa aagctatctg accactgcat cgcggcttta taaatactcg    120
tttgacgaac agcttctgat ttacgcccag cggcggacg ccaccgcctg cgccagcatg    180
gaattatgga atgaggatat gcggcgctgg gtcaaggccg gctccaaggg catcgccctg    240
atcaaaaagg acggcatccg cccacggctg acctatgtgt tcgatgtggc ggataccogc    300
ccggtgcggg gcgcgaaaat gccctatctg tgggaaatga aagacggaca ccatgccgcc    360
gtcatggata cactggcacg gcgttacggg gaaacggaaa agaaggacat cggatctgcg    420
ctgatggagc aggcgaaaca cgccgtagag gaagtgtacc ggaacacct ttctgacctt    480
gcctatgacg cccaggacag tcttttggag gggctggacg atttcaacct ggaagtgcgc    540
ttccgaaatc tcctgaccgc cagcgtccag tatacgctcc tttcccgctg cggcttaaac    600
cccgccgatt acctggagga cgaagatttt gacggcatcc atgaatttc caccccgcc    660
gtcctccacc acttaggcag cgccgccagc gaagtgtcca tgaatatgct gttggaggtc    720
aagaaagcaa tccggcaggc ggaaaaaga aaggcacaga accggcagaa aatcccgaaa    780
aatcccgaaa aaccacttgc aaaagaaccg gtgatcggtt atactaaagt caaggaacag    840
tttaacacat taaaacgcga aagcgaagaa aggagcatac aaaatggcag aactggcata    900
caagaggacg ggcgactacc tgattcccga cttggtgatg gacggggagg acgagatgga    960
ggaaatgccg ctgggcaagt acggcaggct gcggcagacc tttcttcagg aacaccacaa    1020
gggacatac acctcgatgc tgctgaccgg gcggctggaa cccaactgc gggagatcga    1080
ccggcaggcg caggaacagg cagaccggat cgtggcggaa ttaaagaaac agaacggcgt    1140
ggacgagggg atgaaagccc gcgaccagat gggatgggtg caggcagtca acctgtttct    1200
cgcccaggcg gaggaaaccg tactgaacga gatcgtttac aagtaaatca ggagaaccag    1260
caggaagcgg cgggcgaaca gtccgccgtt tctgcgtctg aaaagccggc atttacacag    1320
```

```
ttttctttat tccccactgt ggaacagcag attgaaacta ttgctcaggc acagaaaacc 1380
gagcaggaac tgcgcccgc cgtttcctcc ggcaaggtca cggatgcggt cattggccgc 1440
gccctcacaa gcggcggcaa cgagccggac agcatcctgc gtatcgtcgc ttttttccag 1500
aaagacccta cggagcagtc tgccgccgat tttctgcaaa aggaatacgg cagcggcggc 1560
aaggggctta aaattgccgg gcatgagtat tcgctgtggt tcagccatag cggtatccac 1620
atcgcgccgg ggcgctctga cggggcggc tcccttgtca cctggaagca ggcggcggca 1680
cagatccggc agcttttgaa cagcggccag tttgccagcc aggacaaaat cgacgcggcc 1740
cgtgacaatg aatttcggga attgtcggaa aaactgtggt ttctgcggca gaatttcagc 1800
gaggaagcaa aggaaaaggg cttgctgccc accattgacg cgctgtacgg tggtttcccg 1860
gacagcacgg ctaaaattgc cgcgctcctg aaagacccgg cagaacgagc caaaattgca 1920
caggaagtcc ggcagtttgc catctcccat caggagaacc cgaatctgct gcgcttccgc 1980
cagcggcccg cgccctccga actgtttacc cgcctgaccg atatggataa gccggtgacg 2040
gcatttcatg gggtggaggg ctttgcgcct gcccgtggag cttttgtcac cgaggacgag 2100
attacccgga tgttcactgg cggctctggc gtttccgagg ggaaattccg tatttacgcc 2160
tacttatgc aggggcatga tgccaaagag tgcatcgaat ttctgaaaaa cgaatacggc 2220
attggcgggc atggctacat tggctatgat gaatggcacg acggcaaagg catcaagctg 2280
tcccgcgcc acgatttttc cggtggcaat tatgatacgg ttacgctgaa ctggaagcag 2340
gtacaaaagc ggattgccgg actgattaaa acggatcggt atctgaaccg cagggaaaag 2400
gcatatctgc cggaatatga aaaaatgcaa cttgctcgga gtctatatac cttccaatac 2460
tacgacccca acgatgccag caagacaatc ccccacgact gggatgttga cgcggcaaaa 2520
aaggtattcc gcccgctgtt agatgacccg gaacaatgcg cggcgcacta tgaaaaaatg 2580
gtgaaagccc tggcaatggt atccccggat gagcgcgcct attccctgat ggagccgacg 2640
cttcaaaaaa tgggcgcata cctgcggggc gaatactcct tattcacccc tctgcctgat 2700
gccgtcttgc aggaggaacg gcagaaaaag cagaaacaga aacagaagga aaggcaggat 2760
tccgccagac aaacggctga acctgcaacc ggccttgctg ctgcggcaaa agcccttcc 2820
cgcaagaaac agcctgcccc acgggaggat gcagacgggc agcttacttt agatatgttc 2880
gggctttcgt ccgaaccaga acagccagcc gcaactcctg aaccggagcc ggaactttcg 2940
caggcagaaa ccggcgcatc tgaaacggcc gtccctatcg catcaaagcc ggaaccaccg 3000
caggaaacgc cgaacatacc agaaggggct gcgcccataa tggccgaaag ccccgctgac 3060
cgctatgatt taggctatgg ccatatgggg aacggcctga ccgtctggaa ccgtctggag 3120
gaagaacacg ggcactacaa aacagttgcc catatcgccc ctgaccgtac cgtgactttt 3180
tacgacgcg atatgccgga ggaaatcagg gaaaaaattc aaaaggttgc cgctgccact 3240
gaaatgtcca tttctgccac ccaggacacg ccagtatttt ccaccccgcc ccaggagccg 3300
gaacgggtgc agcctggaa ctcggagcca gagccggaaa aggtgcagga cgcgcccgct 3360
gcggtcaatc cagagccgga aaccggggac agcaatattg taccatcccc cgcgcaaaag 3420
gcgggcgcag aacccaccgg cacaggttcg ccgcagaccc aggaagcccc aaaaacagcc 3480
gccgcagacg gtctgaacct tacgcccaac gtggaggaat acctgaacct gaaagcccaa 3540
tacectgaca aactggtagg cgtccgggta gggaaaaacca tgctgttta cggcacagac 3600
gcagaggaag ccgcccctgc gctgggaacg aaaaccatca cccgcgatat tcccgacctg 3660
gggatggtgt ccatcaccgg agccaatggc tggcagtccg tcttaaaaa gctgctggaa 3720
catggaaaaa gcgtggtgct ggcgcggcct gacacggaac gcggcggcga tgcccctat 3780
gaaatcatca aggaacgcag cgcggcggat tatatccctg tcggcatgga actgaccatt 3840
gacggcggc tgatgaagat tgacagcgt gactataacg ccggaaccgt cagcctgcaa 3900
gatatggaac tgcgggctg gtatcccatt ttccgcagtg aaagcatccc ctttgtgcgg 3960
cagtttgtag aggaagtcca gcaggagcat tttacagcgg aacccatgca ggagccggaa 4020
cgccccgaaa cctccgattt agatacggca aaacagctaa ttgaacagtt gcctacgcc 4080
gaatatggca gcgatgatgt ggattttcc gatttggagc atatcgggat tgcctatggc 4140
acgaccgagg acggcggttt agaagtccag gtggacgtaa acctgctgga tttctccatc 4200
agccagtccg tagacgggaa atgcgtggaa acacggcagt atggcagcct gcgggaactg 4260
attgatatgg aacttgcttt ccttgactat gatatgctgg tttccgtgga acgggatatt 4320
gaagaacggt taaaagccga actgaaccag cggatcaggt ggtcagaaat ggcaggtgcc 4380
agagagggcg tggagccaac agagccggaa atttttccag agaaggacgt ttctccagaa 4440
caggaagcac cgccggaact ggtggagatt gacggcgggc agattacgga aaccccggcg 4500
caaagacaga cccgccgccg cgcccaggag gaagtggaca gccgggtatt ccccagcgag 4560
attattttac agccgctccg tctggagccg gaacgccaca acttccgcat caccgatgaa 4620
aatttaggcg caggcggcga aaagaccaag taccagtaca acgtggaagc catccgcaac 4680
ctgaaacaga ttgaagccga aaaccgcctt gccacaccag aagaacaggc cattttatcc 4740
cgttatgtgg gctggggcg catttcccac gcctttgagc caaacgaccc aaaatgggca 4800
aaggaatacg ccgaactgaa agaactgctc accctgggg agtaccagtc tgcacaaagc 4860
acggttctca acgcccacta caccagcccc accgttatcc aggcaatcta caatgcggtg 4920
gaacagatgg actttacgcc ggggacggtt ttggagccgt ctatgggcat cgggaacttt 4980
ttcgggatgc tgcccgaaaa actgggcggca gcaaagttat atggtggga actggacgac 5040
ctgaccgggc gcatcgccag acagctatac caaaaggcgg atataacggt agacggattt 5100
gagcgcaccg accaccggga cgatttcttt gaccttgcgg tgggcaacgt cccctttggc 5160
agctatcagg tgcatgacaa cgcggtacgac cgtcaaaacc tgatgattca cgactatttc 5220
atcaccaaaa cgctggacaa ggtgcgtccc ggcggcattg tggcctttat caccaccaaa 5280
ggcacaatgg acaagaaaa cagcaaggcg cgggaggcgc tggcacaaaa agccgattta 5340
ttgggtgcag tccgcctgcc cagcaatgcg ttcaaggcca acgccggaac cgaagtcaca 5400
acggacatcc tcttttccca gaagcgcgac cgcatccgg aaaaactccc ggaatgggtg 5460
gaggtcgggc agacagagga cggcatcccc ctgaaccgct acttttaga ccaccccgaa 5520
atggtgctgg ggacaatgac aatgggacgc agtatgtacg gcaatgaaac tgaaaccgcc 5580
```

```
tgccagccca tccccggcgc agaccttcc gggcagcttg cagaagctat ccggcatatc   5640
gcaccgcctg accgggaact gttagaagtg gattccgggc aagacggcga ggaactggaa   5700
agcatcccg ccgaccccac tgtacgcaat ttcagttaca ccctatccaa tgataagctg    5760
tacttccggg aaaactcccg catgaccag gccgtccttg gcaaaacccc cacggagcgc     5820
gtcaggggca tgatcggcat ccgggacgac gcccgccgcc tgatcgactt acagcttgcc    5880
ggcgcagacg acacagaaat ccagtccgag caggcaaaat taaaccggct ctatgatgcg    5940
ttcagcgcaa agtacgggct tcttaacagc actggcaaca agctggcctt tgagcaggat    6000
tcctcctatc cgctgctctg ctctcttgaa atcatcaacg aggaggggaa gctggagcgt     6060
aaagcggata tgttcaccag cgcgcaccatc cagacccacc gcgccgttac cagcgtggac    6120
acggctgtgg aggcgctggc ggtatctatc ggtgaaaaag cctgcgtgga tttgggctat    6180
atggcatccc tcatgggcgg ccctgaaaaa atcccgcaga ttgtggagga cttaaaaggc     6240
atcatcttca aaaacccgc caccggcccc ttgagctggc aggacgggcgg cgcacactgg     6300
catcagggct ggcagacggc ggacgaatac cttccggca acgtccgggc caagttagcg·      6360
attgcccgcg ccgccgctga ggaaaaccg cagtttgcca tcaatgccga aaagctggaa      6420
caggtgcagc ccaaagacct gacggcatct gaaatcagcg tccgtatcgg cgcgtcttgg     6480
attgacccgg agtattacca gcaattcatg tttgaactgc tgcatacacc tgtgtattta      6540
caggacagaa aaatcaagct gcaatattcc cctaccaccg gcgaatggaa cgtgcaggga     6600
aaaagcacag acaacgggga taacgtgcgc gtctatgcta cttacggcac caagcgcatc    6660
aacgcttatg aaatctttga gcagaccttaa aaccagcggg atgtacggat· atttgataag    6720
acagaagtgg acggaaagga ggtgcgcgtc ctcaacgaaa agcagaccgc cattgcccag    6780
cagaagcagg aagccatgtg·cgaaacgttc aaggactgga ttttcaaaga cccggaacgt     6840
cgggaaaccc tttgcaggcg gtacaacgag aaattcaact gcatccgtcc acgcgaatat      6900
gacggttccc atatccgttt tgctgggatg aacccggaaa tctccctgcg tacccaccaa     6960
gagaacgccg tggcgcggat gctctatggg aaaaactcgc tgctgccca ctgtgtcggc      7020
gcgggcaaga cctttgaaat gattgccgcc gcaatggcgcc gcaagcggct gggtctgaac    7080
cagaaatccc tattcgtcgt tcccaaccac ctgacggaac aatgggggcgg cgattttttg      7140
cgcctgtacc ccggcgcaaa ggtgctggtt gccacgaaac gggactttga·gcctgccgc        7200
cgcaagaaat tctgcgcccg gatcgccacc ggcgactatg acgcgataat·tatcggacac       7260
tcacagtttg aaaaaatccc cctctccccg gaacggcaga aagccgtcat taaggggcag      7320
attgatgaaa tcgtcaccgc cattgccgaa gcaaaggcgg aggacggcga gcgctacacc       7380
atcaagcaga tggagaaac aaagaaaaac ctggaagcca agctgcacaa·gctggacggac        7440
ggaaagaaaa aagacagtgt tgttaccttt gaggaattgg gcgttgaccg tctgtttgtg       7500
gacgaggcgc acggctttaa gaacttgttt·ttacacacaa aaatgcggaa tgtggcgggg     7560
attgcccaga cggacgcgca gaaatccagc gatatgtttg ccaagtgccg ctacatggac     7620
gaaatcaccg gcgggcgcgg catcgtgttc gctaccggaa ccccggtgag taactcaatg      7680
gtagaattgt ataccatgat gcgctacctc cagtttgaca·ccctggagca gaacggccac      7740
cgccatttcg acgcatgggc ggcagacttt ggcgaaaagg tcacggcgat·ggaattaaag       7800
ccggagggca gcggcttccg ctccaagacc cggtttgcga aattctataa cctgccggaa      7860
cttatttcta tttggaaaga·agccgccgac·atccagaccg ccgatatgct gaatctcccc·      7920
gtgccaaagg cggaatatat caccgtcacc acggaacca gcggattcca gaaagaaatg      7980
gtggaggaac ttggcgagcg tgcggaaagt gtgcgcggcg ggcaggttga ccccccatatc      8040
gacaaatatgc tgcggattac ttcggacggg cgcaagctgc cgctggatca gcggctgcaa       8100
aatccgctcc tgcccgacga·cccggacagc aaggtcaacg cctgacgtgaa aaatatcgtg      8160
caggaatggc aggccagcac ggagatttta ggaacacaat tagtgtttg cgacctcagc       8220
acacccaagg gggacggctc attcaatgtc tatgatgaca tcaaggaaaa gctgattgcc        8280
aaaggcatcc ccgaagcgga aatcgccttt atccacgacg ccaacacgga aacccaaaag      8340
gcagaactgt tcgccaaggt gcggcggggg caggtgcgtg tcttaatcgg ctccacggca        8400
aaaatgggcg caggcacgaa tgtgcagaac cgcattgtcg cgtcccatga cctggattgc       8460
ccctggcgtc ctgccgattt ggaacacgca gcgggccgcca gggcaacatg                 8520
aatgcgtcgg tgaaaatgtt caagtatgtc accaagggta cattcgacgc·ttacaactgc       8580
gggctggtgg aaaacaaaca gaaattcatc ggtcagatca tgacctctaa·aagccctgcc        8640
cgctctgccg aggatgtgga cgccactgcc ctttcctatg ctgaggtcaa ggcccttgcc       8700
acggcgacg accgtatccg cgaaaaaatg gatttggatg tgcaggtttc caagctgaaa       8760
atgctcaaag ccaaccacac ctcccaacag tacgaaatgc aggacaaagc gctgaaatac       8820
taccgcagga aatcgcgga aaccaagctg ttaatcgaag cgctgacggc agatttgcca        8880
accgtgcagg cccaccccgg t caaggacgat gcgttttcca tgactgtcat ggggcagacc      8940
tacacggagc gcaaggcggc gggcgaagcg gtgattaaag cgtgtatgct gatggacgac       9000
ccggaaaaaa ccgtggattt gggcgagttt cgcggctttc ccatgcagct aaggtgtgac       9060
ggcagcaaat ccgcgtcac catgaagcag aacctgacct attccgccga actgtctgac          9120
gatgcggtgg gcaacgtcac·ccgtattaac aatgcgctga aagcctgac ggagcggttg          9180
gacgcacaaa aagcacggct tgccacactg gaaagcgaac tggaaaacgc caaagaagaa        9240
gccgaccggc cattcccaaa agaagaagaa ctgcgggaaa aatccgcccg·cttaaaccag        9300
cttaaccggg aactgtccag gccggacaaa aaggaggatg agcaggaggt tgaagaaagc        9360
cctgatttgg aggacgccgg aaaaccgcg ccggttgtgc·cgatgcccgc tgtggcggct          9420
ggctataaac cctctatccg tcaggcaatc cgtaactatg aaccgcccac gccggtacat         9480
tccggcatgg agcgcacccc gcgaaaggag gcggtgctg                                9519
```

Seq ID 54
```
atgcctgtgt tgacttcggg gcagcaagaa·aggaaaaagca aacagagaaa aatccgcaac          60
tcagagtatt acgacatgga aggtactttt gacaggctgt atgcagaaag taagaaagat          120
aaaaccttca accatctgat ggaaatcatc gagagtgaag aaaatatcaa gcttgcatac          180
```

```
agaacgataa agaaaaatac aggaagtgat acttcgggag tggacaaaag gacgatagcc    240
gaccttgcaa agttaagcga ggaagaatac gtccgtctca tacggaaaca attcagtaat    300
tatcatccag gacccgtaag gcgagtggag ataccaaagc ccaatggaaa gaccagacct    360
ctgggaattc caactatcgt ggatcggatc gtacagcaat gtattttgca ggtcatggag    420
ccgatatgtg aagcaaaatt cagtgagaac tccaaccgat tccgcccgaa ccgttccgct    480
gagactgcta ttgcacaatg catgaggctg atacaggtac agcatccgta ccatgtggtt    540
gacctcgata ttaaaggctt ctttgacaat atcagccata caaagctgat tcgtcagata    600
tgggcattgg gaatccggga caaaaagctg ttatgcatta taaagagat gctgaaagca    660
cctgttgttc tgccaaatgg agagaaaaca tatcccgcac ggggtacccc gcaaggcggt    720
attttgtcgc cattgcttgc aaacatcgta ttgaatgaac tggattggtg gattgcctca    780
caatgggagg agatgccaac aaaaacaaaa ttcaaaacaa ggagcaatgc gcagggaaca    840
gagattaaga gccacgccta cagggctctg cgccggagca gattaaaaga gatgcatgca    900
gttcggtacg cagatgattt taaaattttc tgcgcaacac atgaggatgc tgtcagagca    960
tataaggcaa cagagttatg gctaaaggac aggctggggc tggaaatcag ccctgacaaa   1020
tccaaagtag ttaatctcaa aagacagtat tctgattttc tcgggttcaa gctgaaagtc   1080
cgtaaaaagg gtaagaaata tgtggttcgg tctcatatga gcgataaagc atacaaaaag   1140
gcgcatgaaa aagtctcgga agaagtaaag aaactgcctt attcgtcaga tgataatgct   1200
cagttcatgc agcttcagaa atataactct gtagttgccg gacttcatga atactattgt   1260
atttctacag aagtatctca tgacttcagc agacttgcct tcagtatcaa caaacagctc   1320
agaaacagac tgaaaggcga cttatccaaa aagggacagc tcagaaatgg gttcattaaa   1380
gaaaagtatg gagcaagcag acagatgagg ttccttcatg gacgtccggt ggttccactg   1440
ggatatgtcc agtcaaagaa tgcccagcat aaaaggaaat ccatcaataa atatacggtc   1500
aaaggtcggg aacagattca taagaatctt gcaatcgaca cagctacaat gttatggctg   1560
atgagaaatc ctgtaaaggg cagaaccatt gaatatgcgg ataaccgtat ttctttgtat   1620
gcggcacagt atggtaaatg tgcagtgaca ggaattccga tgaattccca cgatattcac   1680
tgccaccata aagtaccggt aagcaatggc ggttcggatg aatacgcaaa tcttattctt   1740
gtcagcaaag cggttcatat ccttatccat gcttcttcag aacctacgat tgagaagtac   1800
ctgaaatcct taaatcttga caataaacaa atagaaaagc ttaacaaact tcgttctatg   1860
gcagaaatgc cacctattat tctt                                          1884


Seq ID 55
gtgacagaaa aagcacgcga tttatttgat aaactgatga cgcaaattca gttagatgac     60
acagcgaaag aacatccctt aattcaaaat ggccggattg acaaggtcat tgttcaccaa    120
caaagtcgct tatgggaatt tcatttaagt tttgatgaat tgatgccagt gatgctatat    180
caaacacttt tgcaacaatt ggaactagct tttcaacaaa ttgcgcaggt ttccgttcaa    240
ataccacaa atcagacaac atttacagag gaacaattaa cggattattg gcaattggcc    300
ttgttaaata gtcaatgcaa tacgccgtta gttcagaaag tcttaaaaac acagacacca    360
caatttgaag atcggaaaat tatcttacct gttgataatg aagcagttat tccttatttg    420
aagcaacaat atttaccaat tattgaggaa ctttatttct cttacgggtt tcctaaattt    480
catattgaac caaaaatgga tcaacagcaa gctgcagaag tgctgaaaaa atttgaagag    540
caaaaattag aacaagccgc agcctttcaa caacaagctg ctgaatcgct tgttaaacat    600
gaacaaatga aaaagaaaa acaacaacaa gcgcctgcgt ttgatggtcc aattcgttta    660
ggtcggaata ttcccaatga tgaacccatt atgccaatgg gaaacatgg ggaagaagaa    720
cgtcggataa cgattgaagg ctttattttt gataaagaag tgcgtgaatt gcgttctaaa    780
cgaaaaattt taattttgaa aattactgat tatacttcct ctttcgtggt gaagaaattt    840
tcgaatgggg aaaaagatga acaagttttt gatgcgattc aaccgcaaag ttggattcgt    900
gttcgtggta gtgtacaaga agataccttc atgcgtgatt tagtcatgaa tgctcaagat    960
ttgatggaag tcgcccatgc gccaagaaaa gactacgcgc cagaaggtga aaaacgtgta   1020
gaattacaca tgcatagtaa catgagtaca atggatgcga ccaataaagt cggtgacttg   1080
gtggcgcaag cggggaaatg gggtcataag gccatcgcaa ttacggatca tggcggcgcc   1140
caagcgtttc ctgatgccca tgctgcaggt aaaaaagcag gcgttaaaat tttatacggc   1200
gtggaagcca atattgttga tgatggtgta cccattgctt acaacgaaga acatattgaa   1260
ttaacagatg ctacctatgt ggtctttgac gtggaaacaa cggggctttc cgctgtttat   1320
gatacaatta ttgaattggc cgctgtcaaa atgcacaaag gaaatgtcat tgagactttt   1380
gagcaattta ttgatcctga ccatccaatt tctcaaacaa cgatccattt aactgggatt   1440
acagatgaaa tggtgcgcgg ttcaaaatcc gaagaagaag ttttacgcat gtttaaagag   1500
ttttcagaag gaactatttt agtggcccac aacgcttctt ttgatatggg gttcttaaat   1560
acaagttatg gcaaacatgg aatacctgaa gcggctaatc ccgtcatcga tacattagaa   1620
ttatccagat tcttgtatcc acatttcaaa agccatcgtc taaatacctt atcgaaaaaa   1680
tttggcgtca acttagaaca acatcaccgc gccatctatg actcggaatc aacaggtcat   1740
ttgtgctgga tttttcttaa agaagcgaaa gagaatcatg atatggattt ccatgatgat   1800
ttgaatcgtc atattggcga aggagactca tacaaacgag ctcgaccttt ccatgcaacg   1860
attttagcaa caactcaagc aggattaaaa aatctttta aattgatttc catgtccaat   1920
gtagattatt ttttccgagt accacggatt cctcgttcgc aattatcgaa gttaagagaa   1980
ggacttttaa tcggttccgc atgttcaaac ggcgaaattt ttgaagcaat gatgcaaaaa   2040
ggtgtcgaag aagcaaaaaa tcgtgccaaa ttctacgatt atattgaagt tatgcccaaa   2100
ccagtttacg cgccattaat tgaacaagaa ttggtgaaaa atgaagcaga tttggaagaa   2160
attattagta atctggtgaa aatcggtgac gaacttggta agctggttgt cgcgcactgga   2220
aatgttcatt atttaaatga agaagatgcc atctatcgga aattttagt cgggtctatg   2280
ggcggcgcta atccgttaaa tcgccatagt ttaccaaagg ttcatttcag aaccaccgat   2340
gaaatgttaa cggaatttca atttttagga caagacatcg ctaagcgaat tgtagtagaa   2400
```

```
aatcccaaca aagtcgctga tttatgtgaa gaagttattc cagtaaaaga cgatttgtat   2460
acgccgaaaa ttcctgggtc tgaacaagag atcactgatt taagctacaa tcgtgcgcgc   2520
gaattgtatg gcgatccttt accagaaatt gtcgaaaaac ggttagaaaa ggaactgaat   2580
tcgattaacg ggaatggttt ctcggtcatt tatttaattt ctcaaaaact ggtgcataaa   2640
agtaatgaag atggctattt agttggttca cggggctctg tcggttcgag ttttgttgcg   2700
acaatgacag ggattactga agtaaatcct ttagcgccac attattattg tcctgagtgt   2760
caatattctg agttctatga agatggttcc tatggttcag ggtttgatat gcctgaaaaa   2820
gcctgtccaa aatgtggcgc tcgattgttt aaagatggtc acgatattcc ttttgaaacg   2880
ttcttagggt tccatggcga caaagttccc gatatcgatt tgaacttttc tggcgattat   2940
caagccgaag cccataacta cacgcaaagt ctttttggtg aagaatatgt ctaccgggca   3000
ggaaccattg gtacagtagc cgacaaaaca gcgtatggct ttgttaaagg atatgaaaga   3060
gatcataatt tacatttgcg tggtgctgaa attgatcgct tagccaaagg ttcaacaggg   3120
gttaaacgga cgactgggca gcatcccggc ggtattattg taattcctga ttacatggat   3180
gtttatgact ttacgccaat tcaatacccc gctgatgatc aagaggcaga atggaaaacg   3240
acgcactttg atttccattc cattcacgat aatattttaa aacttgatat tctaggacat   3300
gatgatccga ccgtgattcg gatgttacaa gatttatctg ggattgagcc aaaaaccatc   3360
ccaacagacg atccagaagt gatgcgtatt tttgaaatga ctgatgttct aggtgtggat   3420
gcagagcaaa tttattctaa aacaggaact ttaggaattc ctgaatttgg gacacgtttc   3480
gttcgaggta tgttagaaca gacccatcca tcaactttg cagaattact tcaaatttct   3540
ggactgtctc acggaaccga tgtgtggcta gggaatgcag aagaattgat tcgtcgtggc   3600
gaagccaatc ttgctgaagt aattggttgt cgggatgata tcatggttta tttaattcat   3660
gctggtttgt acagtgggat ggcctttaag attatggaaa cggtacgtaa aggacaatgg   3720
aataagattc ctgatgaact acgggatacg tatttaaatg cgatgaaaga aaacaacgtc   3780
cctgactggt atattgattc ttgttctaaa attaaatata tgttcccgaa agcccatgcg   3840
gcagcatacg tattaatggc gttgcgggtc gcttacttta aagtttattt cccaatctta   3900
tattactgtg cgtttttctc cgttcgggcg gatgacttcg acttagtggc aatgtcacaa   3960
ggaaaagaag ctgtgaaagc acgaatgaaa gaaattacgg acaaagggat ggacgcctcc   4020
actaaagaga aaaacttatt aacggtttta gaattatgta atgaaatgtt ggaacggggc   4080
tataaatttg gcatgattga tttgtacaaa tcagatgctg aaaattttgt gattgaaggg   4140
gataccttga ttgcaccttt ccgagcagtg ccaagtttag gcctcaacgt agcgaaagcc   4200
attgttgctg cgcgagaaga gcaaccgttc ttatcgaaag aagatttggc aacacgaggc   4260
aaagtttcta aaacgttaat tgaatacatg aatgaaaatg tgtgtcttaa agacttgcca   4320
gatgaaaacc aattatcatt atttgatatg tta                                4353
```

Seq ID 56

```
atggaacgaa gcaatcgtaa taaaaaatct tcaaaaaaac cacttattct tggtgttttct   60
gccttggttc taatcgctgc tgccggtggc gggtattatg cttatagtca atggcaagcc   120
aaacaagaat tagccgaagc gaagaaaaca gctactacat ttttaaacgt attgtcaaaa   180
caggaatttg ataagttacc gtccgttgtt caagaagcta gcttaaagaa aaatggctat   240
gatactaaat ctgttgttga aaaataccaa gcaatttatt cagggattca agcagaagga   300
gtcaaagcta gtgatgttca agtcaaaaag gcgaaagaca atcaatacac atttacctat   360
aaattatcga tgacgacgcc tttaggcgaa atgaaagatt tgtcttatca atcaagtatg   420
gccaaaaaag gcgataccta ccaaatcgct tggaagccat cttttaattt tccagatatg   480
tcaggaaatg ataaaatttc gattcaagta gataatgcca aacgtggaga aattgtcgat   540
cgtaatggta gtgggctagc aattaacaaa gtgtttgacg aagtgggcgt agtgcctggc   600
aaactcggtt ctggcgcaga aaaaacagcc aatatcaaag cttttagtga taaattcggc   660
gtttctgttg atgaaatcaa tcaaaagtta agccaggat gggtccaagc agactccttt   720
gtcaatca cagtcgcttc tgaaccagtg acagaattac caacagggc tgcgacaaaa   780
gatacagagt cacgttatta tccgctgggg gaagcagctg cgcaattaat cgggtatacg   840
ggcaccatta ctgcagaaga tattgaaaaa aatccagagt taagtagtac gggcgtaatt   900
ggtaaaactg gcttagaacg agcgtttgat aaagaattaa gaggcaaga tggtggctca   960
ctagttattt tagatgacaa agagaatgtc aaaaaagctt tacaaccaa agagaaaaaa   1020
gatggtcaaa cgattaaatt aacgattgac agtggcgtac aacaacagc ttttgctata   1080
tttgacaaac gccctggtag tgcggtcatt accgatcctc aaaaaggcga tttattagca   1140
acggttagtc caccttcgta tgatcctaat aaaatggcca atggcatttc tcaaaaagaa   1200
tatgatgctt acaataacaa caaagattta ccattcacag cacgttttgc gacaggctat   1260
gcaccaggct ctacctttaa aacaattacc ggtgccattg gttagatgc agggacattg   1320
aagccagatg aagagctaga aattaatggt ttgaaatggc aaaaagataa atcttgggc   1380
ggctattttg cgacacgtgt aaaagaagca agtccagtca atctgcgtac cgctttagtt   1440
aattcagata atatttattt tgcccaacaa acattacgga tgggagaaga caaatttaga   1500
gcgggtctta ataaattcat tttcggtgag gaactagagt taccaattgc catgacgcct   1560
gctcaaattt caaatgaaga taaatttaat tctgaaattc ttttagcaga tacgggctac   1620
ggtcaaggtc aattgttaat ttcgccaatc cagcaagcta ccatgtatag tgttttccaa   1680
aataatggaa cattagtcta tccaaaacta gttttagaca aagaaacgaa gaaaaaagac   1740
aatgtaatca gcgccaacgc agcgaataca attgccacag atctttttagg tagtgtggaa   1800
gatccttcag gttatgttta taatatgtac aatccgaact tctcattggc tgctaaaaca   1860
ggaacagctg aaataaaaga taaacaagat acagatggca agaaaaacag tttccttcta   1920
acgttagatc gtagtaacaa taaattctta acaatgatta tggttgaaaa ttcaggagaa   1980
aatggttctg caacagatat cagtaaacca ttaattgatt acttagaagc aaccattaaa   2040
```

Seq ID 57

```
atgaaaaag tgttaatggg agttcttttcc ctagggttat tattaggtgc agcaacaggt    60
tgtacaagcg atcaagaaaa agccgctggt aaaacaaaag cgtcttctga aaaaacagaa   120
gcaacatctg gtgcctctgc taatggttat acagatccaa gtgaactaaa agacagttat   180
gatgttgtta ttgttggttc aggtggtgcg gggatgaccg ctgcattaca agccaaagaa   240
gctggcatga atccagttat tttagaaaaa atgcctgttg caggtgggaa tacaattaaa   300
tcgtcttcag gaatgaatgc ttcacaaaca aaattccaag aaaaagaagg cattaaagat   360
agcaacgata aattctttga agaaacacta aaaggcggta aaggaaccaa cgatcaagaa   420
ttactacgtt attttgttga tcattcagct gaagcaattg attggttaga tacaaaagga   480
attacattga gcaattttaac gattacaggt ggaatgagcg aaaaacgcac acaccgacct   540
gctgacggat cagcaattgg gggctatcta gttgacggtt tagtacgtaa cgttcgcgaa   600
gaaaaaattc cattgtttgt tgatgctgac gtcacagact tagttgaaga aaacggtcaa   660
atcgatggcg taaaagtgaa aatgaaagat gataaagaaa aaacagttaa agcaaaagca   720
gtcgttgtca caacaggtgg tttcggtgcc aacgaaaaat taattacaca atacaaacca   780
gaattgaaaa actatgttac aacgaaccaa gaaggcacaa caggtgacgg tatccaaatg   840
attcaaaaag ttggcggtgc gttagtggat atgaaagaaa ttcaaattca cccaactgtt   900
caacaaagcg atgctttctt aattggtgaa gcagttcgtg gagaaggcgc aatcttggct   960
tctcaaaaag gagaacgttt tgttaatgag ttagatacac gcgacaaagt ttctgcagca  1020
atcaatgcct taccagaaaa atcagcttat ttagtctttg accaaggcgt tcgcgaccgt  1080
gcaaaagcca ttgatttcta tgatcaaaaa ggtttcgtgg aaaaaggtga aacgattgaa  1140
gaattagctg aaaaaaattgg catgcctgcc gatacattga aagcaacgat tgatacttgg  1200
aaccaagatg ttaatgccaa agatgacaaa caatttggcc ggacaactgg aatggaagct  1260
gacttaagca cagcccctta ctacgcaatt aaaattgcac cagggattca ccacacaatg  1320
ggtggcgtga aaatcaacac gaaaacagaa gttttacgtg aagatggcac accaattaaa  1380
ggtctatatg ctgctggtga attaacaggt ggcttacatg gtcaaaaccg tatcggaggc  1440
aatgcaatcg ccgatattat catttatggt cgtcaagcag gtacgcaatc agcagaattt  1500
gcctctgctc aaaaa                                                   1515
```

Seq ID 58

```
gtgattgttc gcggaggggg cgatctagcg acaggtgtta tccaaaaatt atggcacgtt    60
ggctttaaaa tactagtatt agaaacagaa tgtcctctgg ccattcggcg gactgtttca   120
gtttgtgatg ctattttttca gaaagaacaa agagtggaag atttggttgc tgttcgaatt   180
acttcattaa aggattgcgc gaaatgttgg cagcaaaata agctgcctgt ttttgtcgat   240
caaacagcaa gtgctattca acaattgaaa ccactgattg ttatcgatgc gatttttggct   300
aagaaaaatt taggtacgca tcgaggaatg gcaccgatta cgattgcttt aggaccaggc   360
ttttcagcac cgcaagatgt cgatgttgtg attgagacaa tgcgtggcca taggttgggga   420
cgtttatact ttgaaggaac agctttacct aatacaggta ttccaggtga aattggcgga   480
aaaagtgcgg aacgagttgt ccatgcccct gcttcaggac aagttacccca ccttaaaaac   540
ataggggatc tcgtgttgaa aggggaagca ttattttttaa tcgatcaggt tccagtttat   600
tcgcctctga cagggacttt gagaggattg atttcagaaa aggtaacatg ttatcaaggg   660
ttgaagtgtg cagatgtaga tcctcgacca gttgagaaag tcgattgctt gacgatctcc   720
gacaaagctc gtgcattagg gggagcggtt ttagaagcaa tttttatgat tggaagaagg   780
aaaaacgttt ta                                                      792
```

Seq ID 59

```
atgtattcaa ttattgtaaa tggtcaaagt gaaacctgtg aaacaaacaa aaaattaatg    60
gactttttgc gggaggatct gggcctaact ggaacaaaag atggctgtaa ccaaggctct   120
tgcggtgcat gtacggtgtt agtaaatgga aaggcttcca agcatgtttt gtttactcta   180
gaaaaattag ctggtaaaga agtcacgaca atcgaaggtt taagccaacg gcaaaaagat   240
gtctatgcct atgcgtttgc aaaaacgggt gcagttcagt gcgggtattg catccctggg   300
atggttattt cggcacaagg attattgaac aagaacccg aacccacaaa agaagacatt   360
cagaaagcta tacgaggaaa catttgtcgt tgtacaggat atgtgaaaat aatagaagcc   420
attcagttag ctgctaaaat gttttgtgaa gaagcgagca ttccagaaga acattcaaat   480
ggtaaattag gggaagactt ccaacgggtc gatgcagtgg agaaaacctt aggaacaggg   540
atttatgtag atgacattga catagaagga atgcttcatg catctgcttt gcggagtgcc   600
tatcctcgcg caaaagttct atcaatcgat agtacccaag cattggcgca tcctgactgt   660
gtggctgtgt tcacagctaa agacgttcca ggaaataata aaattggtca tttagaattt   720
atttccgatt gggatgtaat gattccagaa ggagaaataa cacggtatgt tggggatgca   780
gtagcactcg ttgtttcaaa aagaaaagaa accttaccag agataaaaaa tttagtggaa   840
gttgattacg aagaaaatgat accgcttact tcttgtgaag ctgctttagc agaaggtgca   900
ccagctattc atgaaaaggg caatatcttcg tctcatgacg acttggtacg tggcatggaa   960
gaggaagtgt tagaaaattc tgcgtttgta gtgacagaac attattccgt accaattaat  1020
gaacatgcct ttatggagcc cgaatgtgcg attgcacaac ctgaaggtga gggcatttta  1080
ctcttttagtg cgggtcaaag tatctatgat gaacaacggg aagtggctcg gatgctaggg  1140
ctagacaagg aaaaaattca tgttcaatca aaacttgttg gtggtggttt tggtggaaaa  1200
gaagacatga gtgttcagca tcatgcatcc ttagccgctt ggttattgaa gaaaccagta  1260
aaagtactgc taagccgtga agaaagtttg atgttacg ccaaagcgca tggcatggaa  1320
atggacttta cgaccggctg tgatgaagaa ggcaatttaa ctgcaatgaa agcagtgatt  1380
tatgcggata ctgggccta tgcttcttta ggaggaccgg tattgcaacg ggcttgtaca  1440
catgcggcag tccctacaa gtatcaaacg attgatgtgg aagggtttgc tgtatatact  1500
aataatccgc cagcaggtgc atttcgcgga tttggtgttt gccaaacggc ttttgcaatt  1560
```

```
gaaagtaatt taaacttgtt agcagaaaaa gtcggactat caccatggga aatacgcttt  1620
aaaaatgccg tcgccccagg agatacttta ccgaatggac aattggtctc gaaaaatgca  1680
gcgttaaaag aagccttgtt agcggttaaa gacgtctatg aacaagcgga agttgctgga  1740
atttctagtt tctttaaaaa tagtggcgtg ggtgtgggc tacctgatac agggcgttgt  1800
attatttctg ttgaagaagg aaaaatacat gtccgcacaa gtgctgcttg tatcggtcaa  1860
gggatggcaa ccgttacgac gcaaattgcc tgtgaaacat taaatttacc accagaaatg  1920
attgttgcgg aagcccctga tacaagaaga acaccgaatt caggaacgac aacggcttcg  1980
aggcagtcac tcttcacagg agaagcaact agaagagcgg caatgcaatt acggtatgag  2040
cttgatatgg gacgtgcttt aagcgatttg gaaggcgaag aattttatgg ggaatattca  2100
gcgaaaacgg atccattaat taatgataaa aaaagtccag ttagtcatgc tggttatgtt  2160
tatgctgctg aagtagctat tctagatgac aaaggaaaag tagctaaatt tgtggctgcg  2220
tatgatatgg ggcaagtggt gaatccacga gctgctgaag gtcaaattga aggtgggatt  2280
gcaatgggca tggggtatgc tttaactgaa aaatttgcct tagaagaagg ctatgttaaa  2340
gcaaaatatg caacgctagg attagtgaac gccactcaag taccaccaat cgaaacaatt  2400
ttggttcatg ctgacaatct tcacgaaggt ttagcttatg gaattaaagg cgtaggtgag  2460
ttagcaacga tccctaccgc accagcttta gcaggcgcat attttgcctt ggatggtcag  2520
ctccgaccaa gtctgccttt agaaaatacc ccttatcaaa agaaaaaaag a           2571
```

Seq ID 60

```
atgagtgata ttatgcatcc catctcaatt gaggctttat taaattggat ttttagtgag   60
tatcaacaag atggcacaat ttttggcatt cgtaaatttt atcatgcgga tcctacaaaa  120
acgatttcgt tatttggaga aaaaatggaa accccttgtg gtccagcagc aggtcctcat  180
acacaattag cacaaaatat tattgcggcc tatttaacag gttctcgttt ttttgaagtt  240
aagacagtcc agatactgga tggagaagat ttacctgtga gtaaaccatg tatcgcagca  300
gcagatgaat gttacaacgt tgagtggtcc acagaattgc gcgttcccca agcttatgat  360
gaatacgtca aagcatggtc cgtttttgaag ttacttagta aagagttcga gctaggtgac  420
ccgaatggtt tcattttttaa tatgagtgtt ggctatgatt tggcagggat tcaatcacca  480
aaaattgatc gttatattaa cgaaatgcag aacgcagaag gaacgccgat ttgggcggaa  540
tgccaagcag ctgccaaaaa atatctatct tattttaaaa agtagatga cttatatatc  600
gaagcaatta gtcctaaagt ctgtcattct attacgctgt cgactttaca cggttgtcca   660
tcagacgaaa ttgaacgaat tgcggcctat ttattaagtg aaaaagggct ccattccttt  720
attaaatgta atccaaccat gttaggttat gaatatgcgc gtcaaacgat ggatgagcta  780
gggtttgatt atatggtttt tgacgatcac cattttaaag aagatttaca atttgaagaa  840
gcagtaccga tgttacaacg tctgcaatta ttagctaata gtaaaaatct aagcttcgga  900
gttaaaatta cgaacacttt cccggtaacg attgcagcaa atgaattacc tggtgatgaa  960
atgtacatgt caggacgttc gctatttcct cttagtattt cattggcaca aaaattatca  1020
gaagcatttg atgggaaact acaaattttct tattctggcg gtgcagatat tttcaatagt  1080
aaagaaatat tcgatgctgg gatttggcca atcactatgg caacgactct gttgaaacct  1140
ggtggctatc aacgaatgaa tcaagttgct aatgtattaa gcgcagcaga atatcctcag  1200
atggtccatg tcaacttgga taaattagct caagttgttg agaaagctaa aacacaagcg  1260
cgttaccaaa aatcaattaa attaccagaa agtaccaaaa taagaaaaac agttccgtta  1320
acagattgtt atattgctcc ttgtcgaagt gacggtggtc gtccaatcaa tcaggatatt  1380
cctgcttacc ttcgttatgt cagtgaaggt aattatttaa aggccttgca agtaattgtc  1440
gataagaatc cgttgccttt tattactgga accatttgtg cacatccttg tatgactaaa  1500
tgtacacgtc aattttatga agaatctatt catattcggg aagtgaagtt agaagcagct  1560
gaacatgctt atgatgaatt attgacgaca ctagaaaagc cacagccaaa agaaaatgca  1620
cctaaaacag ctgttgtggg tggaggccct gcagggatct ctgctgggta tttattagca  1680
cgtgaaggca tgcctgttac tgtgtttgaa aaatccgaga caattggggg cgtttgtgga  1740
cagattgtac cagaatttcg aatttctatg gaatctgttc aaaaagacgt tcagttagcg  1800
gaattatatg gagccgaatt ccgtacagga caagaggcgc catcattggc ggaattgaaa  1860
aaccaaggct atacgaatgt tatttatgcc attggtgcat ggaaacatgg ggtccttcgt  1920
ttagaatctg gtcgagcatt gaactcgtta gaatttttaa aagcaaatcg agaaaatcca  1980
acgattaatc cttacggaga acaaattgtc gtcgttggtg gcggaaatac agccatggac  2040
tgtgctcgtg cagccacaac attacctggt gttaaaaaag tctcagttgt ttatcgtcgg  2100
aataaacgca atatgccagc agatgaagaa gaactttatt tagctttaga agatggtgtt  2160
gactttttag aattactttc gccaattaaa catgaaatc aacaattaac ttgtgaaaaa  2220
atggtcctag gtgaacgaga tgcctctgga cgccgtcgac caattggtac aggagaaatg  2280
attgatattc ctgctgatac agtgattgcc gctgttggtg aaaaagtaga tacagaattt  2340
taccaagcat taggcattca cacagacaac tatgggaaag tcgtttcaaa tcaagaaaca  2400
ttggaaacaa atatcccagg tgtctatgtc attggtgacg ctaatttagg accagccaca  2460
attgttgaag caatcgctga tgcaaccaaa gcggccaata atatttgtct tgtccacaat  2520
catcactatg aaaaagataa tttaaatagt gatgtcgcat ttgttagaaa caaacgggt  2580
atccttgtga cagacgagat gtcttgtagt caagctagtc gttgcttaga atgctccaca  2640
atttgtgaat cttgtatgga tgtctgtcca aaccgagcaa atattgttgt gtatgtcgaa  2700
ggaaaaccac aaattgtcca tgttgatcgg atgtgtaacg aatgtggcaa ctgtgaaaca  2760
ttttgcccct tatgcaagtgc cccttacaaa gacaaatta cttattttaa cagtgaagct  2820
gatttttatg acagtactaa ttcaggtttc tatgttcaaa atcctgttga taaaaatttgt  2880
ttactgcgtc tatgggggaac ggtgtcaaca attcaattaa cagatcaagg attagatgta  2940
ccagaagatt taattgcatt aatggtaact atgattgaag aatatgctta ttgtatgcaa  3000
gca                                                                3003
```

Seq ID 61
```
atgacaaaac aaacaattaa agagcaaatt ttacatttca tggaaagcca aaaaaagaaa    60
agcttttcaa tggaagaaat tgcacaaggc ttgaatttag aaaaaagttc agattttaaa   120
attttagtgc aaaccattgc acaaatggaa cgagaaaagt cagttagttc caacaaaaaa   180
ggcaaagtcc tgttaccaat gaaagactta ttaatagaag gaacgtttcg tgcaaatgaa   240
cgaggctttg gttttgtaac catcgatcct gaagaaccag atgtttatat tccgaaagag   300
gcaacgaact ttgcaatgga tggagatacg gttttaatcg acgtgatcca acatgcggat   360
ccttttttcag atcgcggcgc agaaggtaaa gtcaaagaaa ttaaagagcg agcagtgagc   420
caagttgtcg gagaatttgt ggcatatagt gaagaagaaa tggcagaaat gggactgtat   480
ggctacatga ttcccaaaga taagaaattg aatcagtata ctgtatcaat tgcacctgaa   540
gggattaagc cagtagatgg tagcattgtc attgctgaaa ttacctatta tccagatcaa   600
gaatatccaa cgagtatgga aggactagtc aaacaagtga ttggtcataa aaatgatcca   660
ggaatggata ttttatcaat cgtggtggct catggaattc ccacagcatt tcctgatgaa   720
gttttggctg aagcggacca agtaccagaa actattgcag aaagcgattt agtcggtcgt   780
cgggatttac gtgatcagtt gattgtgacg attgatggag aagatgcgaa agatttagat   840
gatgcagtaa cggtacaaaa gttagcaaat ggcaatttct ttttaggggt gcatattgca   900
gatgtttctt attatgtaac tgaaggaagc caattggata tggaagcgta tgaacgtggt   960
acaagtgtct atttgacaga ccgagttgtg ccaatgattc cgcagcgatt atcaaatggg  1020
atttgttcgc taaacccaca tgttccacgt ttaaccatga gttgtgagat ggaaattaca  1080
ccagaaggag aagttatttc gcatgaaatt ttccaaagtg taatccaaac aacggagcga  1140
atgacctata cagcagtcaa tgaaatttta gaagagcaaa aacctgaaac gttagaacgc  1200
tacaaagaac tagttcccat gtttaaagag atgggtgagc tgcaccatat attagaagag  1260
atgcgtatgc gccgtggtgc gatttctttt gaagatcgt aagccaaagt cctagttgat  1320
gaaaatggtc atcccaaaga tattcttta cgcacgacgtg gtgtgggtga acgattaatt  1380
gaatccttta tgttggcagc gaatgaaact gtggcgcgtc attatcatga cttaaaactg  1440
ccatttattt atcggattca tgaacaacca aaagaagaaa aaatgcaacg tttctttgac  1500
tttgctgcag tacttggcat tcttgttaaa ggaacaaaag aaaacatttc accgaaagat  1560
ttacaaaaag tttttagagca agttgaaaat aagccggaag aagtggttat caatacgatg  1620
ttgctaagaa gtatgcaaca agcgaaatac tcagaagaca actacggaca ctatgggttg  1680
gctgctgaat attatacgca tttcacttca ccaatacgtc gttatccaga tttaattgtg  1740
catcgtttga ttcgcagcta tagccaagat caatctgaaa aaaatcaaga aaaatggaac  1800
gaagcattac cagaaattgc caatcatagt tcaagtatgg aacgtcgcgc agttgatgca  1860
gagcgtgaag tggacgccat gaagaaagcc gaatttatgg tagataaagt gggagaaacg  1920
tatgacggga ttatcagctc agtcacaaaa tttggtatct ttgtggaact gcctaataca  1980
atagaaggct tgattcacgt gaacaactta aaacaagatt atttccactt tattgaaaat  2040
catatggcgt tagttggtca acggacggga atgactttga aaattggtca gaaagttcaa  2100
atccgtgttg aaaaagctga tccagaaaca agagaagttg attttgaatt gatttcagct  2160
gaggaagtcg cgccagtcga aggaccaaaa ggacgtaaaa aaggcaaagc caattcttca  2220
actcgttcaa ataatcaacg aagaaataaa aaagatgaat catttgatgg taagaaaaag  2280
aaaaataaga aaaaaggcaa aggcaaaaaa caaccttttt ataaagaagc aatgaaacaa  2340
aaaaataaaa aagggaaaaa gaagaag                                       2367
```

Seq ID 62
```
atgaaaaata gatattttat tctgatgatc ctaagttttt atttctttgc ttttgggatt    60
gaaacctcag cagcagaatt acgatttagt gttgaaacag aaattccaga aaaccaaata   120
gataaaacaa aaacatattt tgattaaatg atgaaaccag accaagaaca aattttgaaa   180
gttcgagcgg ccaatagtac agatgaaaat ttagtgattg atgttagtgt gaagtcggcc   240
acaacaaaca gcaacggagt gattgaatat ggtgagtcgt tgacagcttt agataaatct   300
gcaccagcga atttatcaga aatcattcaa ctgaaagacg gcggtgagag tgttgaattg   360
cctgctaaaa gtgaaaaagc agttgaattg cgagtgaaga tgcctaaaga agagttttca   420
gggcaactag cgggcgggat cacgtttagt gaaaaagtcg atgaaacgaa agacaagcaa   480
aaagaaaata caaacggctt agcaattgaa aatcgctatg cctatacagt agctgtttta   540
cttcgagaaa atgaaacagt tgtacaacca gaattgtctt tagagaaagt tgaaccgaca   600
caaagaaatg cgcgcagtgt tatttctgca acactgttaa accatgaagc agcctatctt   660
caatcaatga agtaacggc aaatgttaaa aataagaaaa caaataacgt tattttagag   720
aaagaacaag aagatatgca aatggcgccc aattcaattt ttaatttccc tattccttat   780
gaagaaaatg aaatggaagc tgggacgtat gtattggcaa tgactgtgga aggctcagga   840
aaaaaatggc aatttacgaa agaatttacg atctcaaaag aggaagcaaa aacatttaat   900
gagaaagatg taaccgtcaa aaaaacggaa tctaaactta tttatctctt aatcggttta   960
ttgatccttt tactgataat ctgcctgttc atcatttac gtctgaaaaa acaaaaaaat  1020
aaa                                                                1023
```

Seq ID 63
```
atgagtagga agcgaaaaat cagcttaatt agtttagtca tcattttggt ttttgtcaca    60
gtcggctcag catactttgc tgtagcgggt agctatttaa agaaaacaat tgataaaggc   120
tatgttccca taaaaaatga ttataatgaa gcgcaaaata agatagtca atcgttttg    180
attatggggc tagacaatac aattgaacgac aaattaggca caactaggac tgatgctatg   240
atggtgatta ccgtgaataa caagacgaag aaaataacct atttaagttt gccacgggat   300
agttttgttc aaattgatgc gaaaaattac caagggatgc agcgaattga agccgcctat   360
acctacgatg gaccaacagc ttctgttaac acagttgaga attattgaa tattccaatc   420
aatcattacg ttgtgtttaa cttttttatct tttattaagt taattgatgc ggttggcggc   480
```

```
atagatgtca atgtcaagca ggcgtttgat ggtgtcacca aagacgggcc aggatccatt    540
cattttgatg cagggaaaca gcatttagat ggtacgaaag ctttatctta tgcccgtgaa    600
agacatagcg ataacgatat tatgcgtgga ttccgacaac aagaaattat tcaagcagtt    660
gaagacaagt tgaaatctgg tcaatcaatc atgaaaataa tggacattat tgattcgtta    720
aatggaaaca ttcaaactga tgtggattcc aatgaattga ctcatttagt caaagaaggt    780
ttgacttgga ccaattatga taaacaacag ctttcttttg actggcgcac ttttagtaat    840
gaagggcgca gtatggttga actatacsca gatagtattg aaaatgtccg tcatcaatta    900
cgtgtgtctt taaatttaga aaagccagat gaacgagatc aagacggcta tgtcttccat    960
acgaacggtg aatttttata tcaaagtgat tataccgttc aagatgaagc agctgaggaa   1020
aacgaaatga cttccatcaa cggcaatacg tatattggtg ttcctggtaa tacacagacc   1080
ggcccgttgc catcagttaa aacggaaaat ggctttataa aa                      1122

Seq ID 64
atgaaaaaaa gtcaagctgt tcaattaatt caggaacttt caaatgctaa tggtgtttca     60
ggctttgaaa cagaagtcgt tagaatcctt caacacgcaa ctgcggattt tactattcaa    120
cggctagatt ccattaaaaa tttatatctt gagagaaaa ataatctaag tgaaggtcca    180
gtagtgcttt ttgatgctca tagtgacgaa gtgggtttta tgattcaagc aattaaagaa    240
aacggcttgt tgcgtttttt tccgttaggt ggctgggtcc ctaatacgat ttctgcacaa    300
aaagttcgga ttcgtaaccg ggaaggtacc tatctgccag gcgtagttac tagtcggccc    360
cctcatttta tgacaccaga ggaacgtcag cgaccgttaa ccattgctga tttgacgatc    420
gacattggtg ctacctcaaa agaagaagta attgaaactt acaaaatcga tcttggtgca    480
ccggtcattc cagatgtgac gtgctgttac aacgaacaga cagatctttt tttaggaaaa    540
gctttttgatt gtcgcattgg ctgtgcttgc ttagtggatg tcatggagga gttgaaagaa    600
gagacgcttc cattcaagtt agttgcgaca gtaactgcac aagaagaagt tggcgaaaga    660
ggagcactta ttgccgcaaa acaagtcaat cctgattag ctattgtttt cgaaggctgc    720
ccggctgatg atacggcaga aacgccagag atgattcaat cagcaatggg aaaaggaccg    780
atgttacgtt attttgatgt ttctatgatt acgaatccgg aatttcaaga gtatgcacta    840
gagattgcta aaatacacaa aattcctgtt caagtttctg ttcgaagtgg tggtggaacc    900
aatggcatgg cgattacaca agttcaagga gccccgacga ttgttgtagg aattcctgtc    960
cgttacgctc acacacccca ttgctatgta gattttcaag attaccaagc ggcgaaagaa   1020
ttagtcattc aattaatcaa aaatttagat gctgacaaaa ttcaagcact ggttcagcca   1080
ttgtcaaagg agtggaacaa a                                            1101

Seq ID 65
atgaccacga gtcaaatgat tgccgatttt accacacttg ctattcaagc aggtggttgg     60
atggaaatcg atcgcttata tcttcagaac cgtttattat caatgcttgg tgaacaagaa    120
ttaggcgaag tagacatacg gcctgtggca acaccgcccg ctgatttggc ggaacaattg    180
tgtcaagttg cgagtgccaa ccaacttgta aaaacagagc aacaaaaga acagttcatg    240
gtgcaattga tggatttgtt aacaccgcca ccttctgtgg ttaatgcttt ttttgcacaa    300
cattatgcca aagagccaca agaggctacc gagtatttt accaactttg tcaaaaaaac    360
ggtacggtca ttgaacagga agaacctgtt gtttttttcaa cagtctatgg agatttttta    420
gcaaataagg tcccacagtga ggcgtcaaaa gcaacgctct ctgcacaaag ctatccccag    480
tgtgaatggt gcatggcaac agaaggctac cagggaagtc agcaatttcc tgcgactacg    540
aaccatcgag tgattcgaat gaacttagat ggcgaaagct gggggttttc ctttgtaaaa    600
caggcacaat atcagcaaca aggtgtgatt gcgtttgaaa aacatcagtc agcgaaacgt    660
tcaattaaga cttttcagca attgttgaaa attgttgagg tctttcctca ttatttgct    720
gggattgatg ctgactttga acaaaatgag catgtctatt atcaaacggg tttgcaacaa    780
tttccattag ccgaagcgag catttcagaa tatgttgagt tagctaacta tcctttaatc    840
aacgcaggaa tggtgaattg gccagttgag acatttcgcc tagagggacc gaatgcgtca    900
gaagttgccc aagccgcaaa tgatattttt gagcaatggc aaatgctaaa actgccaacg    960
gatgagatac aaatcgttgc ccgacgaaaa gaactttat atgtcatgga cttaattttt   1020
agtcgacctc aagcaaagcc ttcgttaacg ttagcagaag tccaaggttt aacgacttgg   1080
aacaaccaga aaacgcaggc acttgaaact gtcgcttcag cgtatcaaca gcgattaaaa   1140
gaagccagtg cgtttgcaga aacaagtgaa ggaaaagcgg ctttttttagc aatggttgcg   1200
ccagtgactc at                                                       1212

Seq ID 66
atgagtattt tatgggtaat caatggcatt ttgctagtta ttgtactagc gatggtattc     60
aatgaattat atttgaaaat tatggtgaaa cgttcagcaa aaatgctgac ggaagaagaa    120
tttaaagaaa caatgcgtaa agcacaagtg attgacgttc gtgagaaaga tacatttgat    180
gctgggcata ttttaggggc gcgcagcatg ccatacagca tgttaaaaac aacaattggt    240
tctcttcgta aggatcaacc tgtttatctt tatgatcaga agaaagcttt aagtatccga    300
gcagctaatc tattacgcaa aaatggctat actgatattt atattttaaa aggcggctat    360
gatggctgga ctgggaaagt gaaaaagaga aatagt                            396

Seq ID 67
gtgattttta tgaatcctga tgaaccaaat ttcatttgga aagatttgaa tgcagttcgt     60
gatatggggt gcattatcga aaatgagctg tcagaggttt taccaaataa acgatatgaa    120
acgtattcga ttatcggaag aagtggtgaa tttaatgaaa cgttcaatga ttatgaaccc    180
tttgattatg aaattgaaga tgtaactatt ccatatgaaa atttagcggc agtcaaaaga    240
tggttaactg gtaaaagtaa acttattact cacaatgatg aagataaaata tttagatgct    300
```

```
atttgtacaa tgagtaaacc aacttcattc aaaaatgaat ggggtgtttt ttataccttt   360
aacgttgaat ttagatgtca accgttcaaa agaaaagtaa acgaacaacc aaaagtgatt   420
aaaacaaaat caattgaaat tactgatcac ggtgatgaaa ttgcttttcc ttatatcgaa   480
attaattcaa aaggtggcga tattacgtta aacattggta gtaactcact aacgattttg   540
cgtacacaat caggaatcgt cactattgat accgaaaagg gaaaagcaat acaagaagga   600
aatccactat ttacacgcgg cagttggata aaaacgaatc ccggtcaaaa taaattaaat   660
atatcaggaa attttatga agctaagttt tggaatagga gcgcgtattt a             711


Seq ID 68
atgtctattc acattacttt tccagatggc gctgttaaac cgtttgattc tggaattaca    60
acatttgatg ttgctaaaag tattagcaac agtttagcca aaaaagcttt agctggtaaa   120
ttcaatggtg tttttaatcga tttagatcgt cctatcgtag aagatggttc gcttgaaatc   180
gtgacacctg atcatgaaga tgctttagga attttacgtc attcatcagc tcatttaatg   240
gctaatgcct tacgccgtct tttccctaac attaaatttg gcgtaggtcc tgcgattgat   300
tctgggttct attatgatac agataatgga gaatcccctg tgacagcgga agatttacct   360
gcaattgaag ccgaaatgat gaaaattgtg aaggaaaata acccaatcgt tcgtaaagaa   420
atctcacgtg cagaagcgtt agaattattt gctgatgatc cttacaaagt tgaattaatt   480
acagatttgc cagaagatga aatcatcact gtctatgatc aaggcgattt tgttgattta   540
tgtcgtggtg ttcacgtccc ttcaacagga cggattcaag tctttaaatt actttcagta   600
gctggtgctt attggcgcgg aaactctgac aatcatatga tgcaacggat ttatggcact   660
gcctttttg ataaaaaaga tttaaaagag tttatcaaaa tgcgcgagga agccaaagaa   720
cgtgaccacc gtaaattagg aaaagaatta gatttatta tggtttcaca agaagttggt   780
tcagggttac ctttctggtt accaaaaggc gcaaccattc gtcgtacaat tgaacgttat   840
attgtggaca aagaaattag cttaggttac caacatgtgt atacaccaat tatggcagat   900
gtggaattat acaaaacatc tggtcactgg gatcattacc atgaagatat gttcccacca   960
atggatatgg gtgatggcga aatgctggta ttacgtccaa tgaactgtcc acaccatatg  1020
atggtttata aaaatgacat tcatagttac cgcgaattgc caattcgaat cgctgaatta  1080
gggatgatgc accgctatga aaaatctggc gcattatcag ggttacaacg tgttcgtgaa  1140
atgactttaa acgatggcca tacttttgtt cgtcctgacc aaattaaaga cgaatttaaa  1200
cgtactttgg agttaatggt ggcagtctat gctgactttg acattacgga ttatcgttta  1260
cgcttaagct atcgtgatcc aaataataca gacaaatatt ttgatgatga tgcgatgtgg  1320
gaaaaagcgc aagcgatgtt aaaagctgcc atggatgaat tagaattaga ttactttgaa  1380
gcagaaggcg aagctgcctt ttacggtccg aagttagatg ttcaagtaaa aacagcttta  1440
ggaatggaag aaacattatc aaccatccaa ttagacttct tattaccaga acgttttgac  1500
ttaacttatg ttggcgaaga tggtgaaaat acacatcgcc cagttgttat ccaccgtggt  1560
attgtctcaa caatggaacg atttgtggct tacttaacaa aagtttacaa aggcgctttc  1620
cctacttggt tagcaccaat tcaagcaact attatcccag tttctgtaga agcgcattct  1680
gagtatgctt atgaaatcaa agaacgttta caagcacaag gcttacgtgt tgaagtcgat  1740
gatcgtaacg aaaaaatggg ctacaaaatt cgggcctctc aaacacaaaa agtaccttat  1800
caattagtgg tcggggacaa agaaatggaa gacgcaacgg tgaacgtccg tcgttatgga  1860
agcaaagaaa cgtctgtcga agatttatca attttcattg acagcatggc tgctgaagtt  1920
cacaattaca gccgt                                                    1935


Seq ID 69
gtgtctaaca aaaaagtcgc tggaacaatc atactgaact taaacgatgg tagtaaaaaa    60
tttttgatgc atccaatagg tgaagcgatt gagtttgcga tggcaaaagt ttctgatgag   120
atgacgggtt tagctagtat gctacaatgg ttcaaagaag aggttcaatt ggacgttact   180
tcaattagct tagtagaatt aactaacgcg catatcaata aagaaaatgt tccgctgttt   240
gtcttcgaga tggatgaaca agcacttaaa aatgaaatgg ataaagacta tacatgggta   300
gcacctgcag aattaaaaac gatttggtca aaatatcata ttgaaggtgt cccaatgttt   360


Seq ID 70
atgaagaaat ttgatggcgt aattttgat atggatgggt tgcttttga tacggagttg    60
atttactata catctactca aaaagtagca gatgcaatgg ggttaccata tagtaaagaa   120
gtctatttag actatgttgg gatttccgat gaagaagtac aagagaatta ccgacgtatt   180
tatgcatcct atggtcatga cacagtggaa gaatttatcc gccgttctta tgatgatacc   240
ttgcaagagt ttcgatctgg taacgtccct ttaaaaccag gggtggtgga attcctcgac   300
tttttagatg atcaaaaaat tcctcgatta gttgcttcaa gtaacgtccg tccagcaatt   360
gaaatgttat taagccatgc tggaattcag gatcgttttg taggtattgt ttcagcagag   420
gatgttaaac gagcaaaacc ggacccagaa attttcaaa aagctcgtca acttttagga   480
acagaagccc cgaaaacctt gattttcgaa gactcattc atggggtaag tgcggctcat   540
agtgcgggca ttccagtgat catggtacct gatttattcg aaccaactga agttatccaa   600
gagaaaacat tacatgtctt ggagagttta catcaagcgc cacattattt aaaa          654


Seq ID 71
atgaacgaat tagtattttt acattcacag tatatcgagg aagaacctta tacgaccgat    60
gaggttattg caaagtattc acaaattaaa agagaaagtg tttctaaact aataaaaaag   120
tatcaaaaag acctagaaga gtttggcaag gtcggatttg aaatccgagc aatggagagt   180
gggcaaaaag ctaaattcta tcaattaaac gaggaacaag cgaccttgct tatcacttat   240
ctagataata ccgaaccagt tcgacgattt aagaaagcgc tggtgcgaca gttctatgac   300
atgaaaaacg gactctacgc tagacgtatg gaacgacaaa aggaaaaaag tgtgcgtaag   360
```

```
tcaatgacgg atgttatcaa ggagttaggg ctttccccgc attactacaa gcattataca      420
gaccttgttt ataagacagc attaggtttt aatgctaagc aattaagaga agctcgagag      480
gtaagtaaaa aatcaacgat attagatttt ttaacgtcag aagagatcga agcagtgaac      540
aaacgagagc aacaagtagc tacattgctt actcttaaaa tggattatga cacgattaag      600
agcattttga atggtcaagg tgttttatat caaacaacgc ttaaaatgcc agtagcgaca      660
aat                                                                    663
```

Seq ID 72
```
atgccaaaac caggcgaaac aattcacgcc attgaacatt ttacagcagg tggtggtaaa       60
ggagcaaacc aagcagttgc agcgaaacgt tcaggcgcag aaacatattt tattggtgct      120
gtgggaaatg atggcgctgg agctatgatg actgatttaa tgagtcaaga tgaaattaat      180
ttaactgggg tcaccacttt agagaaaaca gctactggcc aagcctttat tatggtagat      240
aacgctggtg aaaatagcat tatgatttac gcaggagcca ataacgcctt tacgccaaaa      300
caagtccagg aacatcaaga aattattgaa aaaagtgatt ttgtgattgc ccaatttgaa      360
agtgcgattg atagtacgat tgcagcgttt aaaattgcta aaaaagcagg cgtcaaaacg      420
atttttaaatc ccgcacctgc gttagaacaa gttcctgaag aattactaaa cgtgacagat      480
atgattgtac caaatgaaac agaaaccgaa attttaacag gcattaaaat cacagatgaa      540
gcgagtatgc gtaaagccgc agaagcactt catcaattag ggattgaagc agtaattatt      600
acagtaggta gtaaaggcgc cttttatgac gtcaatggac gaagtggtat tgtgcctgct      660
tttaaagtga aagcggttga tacaacggct gctggcgata cttttattgg cgcattaagt      720
agtatattag aaaaagattt tagcaatttg gaagaagcta ttcgttatgg aaacaaggcg      780
tcttcgttga ctgttcaacg ttttggagcc caaccttcga ttccttatca acacgaattg      840
gcagacaaa                                                              849
```

Seq ID 73
```
atgattaagc tagcccatat tcataaatat tattattccg aagaagaaac actgcatgtg       60
ttagatgata ttaacctcca agtagacgct ggagaatttt tagcgattat gggtccttct      120
ggttctggga aatcgacttt gattaatttg ctaggcttta ttgataaaaa gtttgaagga      180
acgtatttat ttgaagaccg tgaaattggc gactttctg ataaggaact atcccgaatt      240
agaaatgaag cagtcggctt tgtctttcag aattttagtt taattgaaac actaacagtt      300
gaagaaaaca ttgaataccc tcttttatat agtgggttaa ccccaaaaga agccaaagat      360
cgtgtccacg aagttctgac aaaagtcggc ttgccagata aaggcaagaa gcatccaaaa      420
caactttcgg gcggacaaca acaacgtgtt gccattgcgc gtgcgattgt caatcggcca      480
agtttattta ttgcagatga accaactggg gctttagata gtaaaacatc ggaagaaatt      540
ttaacgctct ttcagcaatt aaacaatgaa ggtgtcacca ttattttagt gacacatgat      600
gaagaaacaa ttgaatactg caatcgtttg attaaagttc gtgatggaaa gattttagag      660
gaggtgctga ca                                                          672
```

Seq ID 74
```
atgaaaaagt caggccagaa gaagcgaaat aaaaaagtgt ggttcggtat tggtgccgca       60
gtggtcgtcg tcggatttat cggagcaaaa acagtttttt cttcaaaaga agttgaacct      120
gaatacacaa cgtataccat tacagaaatg gcctcattaa aattagatgg tcaagtgagt      180
tttctagata ctcgggatat tttctttgat ccttctttga gaaaaattgc cgaaataaat      240
gttgaaaatg gtaaagaagt caagaaagat agtccattgt taacgtacaa taactcagac      300
atccaagcca cagaaacaga acaagcgaat gctgttaatc ggaacaatct tcaagttcaa      360
caagcacaag agaacgttaa cttagccact caaaaataca atgaagcgtt aaacaaagtt      420
gctgcagcca agcaaaaatt aaacaccgca aaagaggctg aagagaaaga aacactaaat      480
gcggaaattc agcaattgaa tgaagccgtt agtgctgcga atagtgaagt cgcccaagct      540
aaccaagccc ttcaattagc caacagtgac gccgtcggag cagcaaacac gcttgaacaa      600
acacgtggca aagtcaacac agtcgttaca gcaccaattg atggccaagt cactgttgat      660
gctagtgcca tgagtagtac cgacaaacca gtcattaaaa ttgcaacaca aaagaaaaat      720
attcaaggaa aagtaaccga gtatgattac gataaattac agacagggga agaagtaaca      780
gtgacgaccg ttggttcagg taaatcagca ccaggcaaaa ttgtttccat tgcgcaaacc      840
ccaatagcta aaaatgaagg caatccagta gtgagttatc aatttacagt agaaggtgat      900
tttccttggg cagaaggctt gtctactagt atcgctgtgc cacaaaagca aatgattata      960
ccgacagcag caatccaaaa agaaggccaa aaagaatttag tctatgtttca caaagcgggc     1020
aaagcgaaaa aaacaccgat tgaaacagaa acaaatttag gtcgtaaagt tgtcaaaagt     1080
ggcttgaatt ggaaagacca agtgattgcc aatcccaata aagaattaaa agataatcaa     1140
gatgttcagg tagctgccaa tgat                                            1164
```

Seq ID 75
```
atgattaaaa aaataatagt cgttgttgct ttcatgctta caggcttctc gctaactgcg       60
atgagtgcat ctgcagaaga aataactgat ttatttttac aaaaagaagt gacatattct      120
ggtgtagaag gaggaaaaat tggagagaat tggaaatacc ctcaatttgt tggcgaaaaa      180
gctgtcgatg gcgatgaaac aacacgctgg tccgctgata gcaagatga acaatggtta       240
attgttgatt tgggtgaagt aaaaaacatt ggcgaactcg tttttacagtt gcatgcagaa      300
agtcctgtgt atgaaatctt agtttcaaca gatggtgaaa gctatcaatc tatctttaaa      360
gaagagaatg ggaaaggtgg ccaacctact aaaaaatata ttgatggcaa taatgtacaa      420
gcacgttttg taaaatatca gcagatgaaa atgtggcaac acacaaataa gcaattttac      480
agttcaagta ttatttcgtt tgaagcatat gaaaaaaaac gactgccaga agcgattaaa      540
cttttaacag agaacctgac tattagtgaa aaaagaaagc aacagctagc ctttgaagtg      600
```

```
tcaccagctg gagtagatat cacagaggat cagattgagt ggtcgagtag tgatcctacc    660
attgtgacgg ttgaccaaac gggtaattta acagcagtta agagtggtga agcgaaggta    720
acagtcaaaa ttaaaggaac agaaattagt gatacaattc ctgtgactgt tgtagcagaa    780
aataaacaat atgcggaaat gcgagcaaaa tggaaaatgc gattgttagg tacaacgcag    840
tatgataacg atgcagatgt gcaacagtat cgtgcccaga ttgctacgga gagtttagca    900
ttatggcaaa cgcttaatca agcagcagat cgtgagtatc tgtgggaacg aaaaccatcg    960
gatacagtgt ctgctgatta cacgactcaa tttaccaata ttaaaaaatt agcgttaggc   1020
tactatgaac catcaagtga gctttttgaa aaaccagaag tttatgatgc gattgttaaa   1080
ggcattgaat ttatgattga tacaaaaaaa tacaatggaa cgtattacac aggtaactgg   1140
tgggattggc aaattggttc cgcgcagccg ctaacagata cattgatttt attacatgat   1200
gacctattga atacagatgc agaaaaatta aataaattta ctgctccgct gatgctgtat   1260
gcaaaagatc caaacataca atggccaatt tatcgtgcaa caggagctaa cttaacagat   1320
atttcaatca ccgttttagg tactggactt ttgttagaag ataatcaacg cctagtacaa   1380
gtacaagaag ctgttccgtc cgtttttaaa agtgtttcct ctggtgatgg cttatatcct   1440
gatggttcct tgattcaaca tggttatttt ccgtacaacg gcagttacgg gaatgagttg   1500
ctaaaagggt ttggacgaat tcagactatt ttacaaggtt ccgactggga gatgaatgac   1560
cctaacatta gtaatttatt taatgttgtg gataaaggtt acttacaatt gatggtaaat   1620
ggaaaaatgc catccgatgg ttctggtaga agtatttcca gagcgccaga aacgaatcct   1680
tttactacag agtttgaatc gggtaaagaa acaatagcta atttaacctt aattgcaaaa   1740
tttgcaccag aaaatttaag aaatgacatt tatacatcta tccaaacgtg gcttcaacaa   1800
agtgggtcat actatcattt cttaaaaaa ccaagagatt ttgaagcgtt aattgacttg   1860
aaaaatgtag tgaatagtgc gtcacctgcc aagcgacac caatgcaatc tttaaatgta   1920
tatgttcga tggatcgagt cctacagaaa aataacgaat atgcggtggg gatcagtatg   1980
tattcacaac gtgtcggaaa ctatgaattt gggaatacgg aaaataaaaa aggctggcat   2040
acagcagacg gcatgcttta tttatacaat caagactttg ctcagtttga tgaaggatac   2100
tgggcaacga tcgatccata tcgattacca ggaacgacag ttgacacaag agaattggca   2160
aatggtgctt atacaggaa acgcagtccc cagtcatggg taggtggctc aaataatgga   2220
caggttgcct ctataggaat gtttttagat aaaagtaatg aaggaatgaa cttagttgct   2280
aaaaaatctt ggttcttatt agatggtcaa atcattaatt tgggaagtgg cattactggt   2340
acgacagatg cttcgattga aacaatcctc gataatcgga tgattcatcc acaggaagtg   2400
aagcttaacc aaggttcaga caaagataat tcttggatta gtttaagcgc agcgaatcca   2460
ttgaataaca ttggctatgt ttttcctaat tcaatgaata cgcttgatgt tcaaatagaa   2520
gaacgctctg tcgctacgg agatattaac gaatactttg ttaatgataa aacctataca   2580
aatacatttg ctaaaattag taaaaattat ggcaagactg ttgaaaatgg tacttacgaa   2640
tatttaacag tggttgggaa aacgaatgaa gaaatcgcag ctctttctaa aaaacaaaggc   2700
tatactgttc tagaaaatac agcaaactta caagccattg aagcaggtaa ttatgtcatg   2760
atgaatacat ggaataatga ccaagaaatt gcaggactgt atgcgtatga tccaatgtcg   2820
gttatttcag aaaaaaattga taacggtgtt tatcgcttaa ctcttgcgaa tcctttacaa   2880
aataatgcat ccgtttctat tgaatttgat aagggcattc ttgaagtagt cgcagcggac   2940
ccagaaattt ctgttgacca aaatatattc acttaaata gtgcggggtt aaatggcagc   3000
tcgcgttcaa tcattgttaa aacaactcct gaagtaacga aagaagcgtt agaaaaatta   3060
attcaggaac aaaaagaaca ccaagaaaaa gactacacgg caagcagctg gaaagtctac   3120
agcgaagcat tgaaacaagc acaaactgtg gcagatcaaa caacagcaac gcaagcagaa   3180
gtagaccaag cagaaacaga gttacgttcg gcagtgaagc aattggtaaa agtgccaact   3240
aaagaagtag ataaaaccaa cttgttgaaa atcatcaaag aaaacgagaa acaccaagaa   3300
aaagactaca ccgcaagcag ctggaaagtc tacagcgaag cattgaaaca agcacaaact   3360
gtggcagatc aaacaacagc gacacaagca gaagtagacc aagcagaagc aaaactacgt   3420
tcggcggtga gcaattggt aaaagtgcca actaaagaag tagataaaac caacttgttg   3480
aaaatcatca agaaaaacga gaaacaccaa gaaaaagact acaccgcaag cagctggaaa   3540
gtctacagcg aagcattgaa acaagcacaa actgtggcag atcaaacaac agcaacgcaa   3600
gcagaagtag accaagcaga aacagagtta cgttcggcag tgaagcaatt ggtaaaagtg   3660
ccaactaaag aagtagataa aaccaacttg ttgaaaatca tcaaagaaaa cgagaaacac   3720
caagaaaaag actacaccgc aagcagttgg aaagtctaca gtgaagcatt gaagcaagcg   3780
caaactgtgt cagatcaaac agcagaacg caagcagaag tagaccaagc agaagcaaaa   3840
ctacgttcgg cagtgaagcg attaacattg aaaaatagtg gggaaaataa aaaggagcaa   3900
aaaaatgggg ggaataatgg acacttaaat actagtacag gagttgatca aactggtacg   3960
aaacaagtta gccatcaag ccaaggtggt ttcagaaaag ctagccaatt tttaccgagc   4020
acaggagaaa agaaatcgat cgcgcttgtg attattggtc ttctagttat cgccagtggg   4080
tgtcttttag tttttcgtaa aagtaaatcg aagaag                             4116

Seq ID 76
atgaagaaat ctgtttttatt tacttcatta cttgtattat caagcttagc tttagcggcc    60
tgcggcggtg gcagtgacga taaaggagct agcaacggcg gcagcgacaa ccaagtatac   120
acaatggttg aatcccaaga aatgcctagt gccgatccgt cccttgcgac agatgaagtg   180
agttttacca ctttaaataa tgtctacgaa ggaatctatc gtttagataa agacaacaaa   240
cccgcgcctg ctggtgcagc cgaaaaagcg actgtttcag aagacggttt agtttacaaa   300
gttaaattac gtgaagaatc aaatatggtct gatggcaaac cagttactgc tgcagattac   360
gtttacggtt ggcaacgaac agtggatcct gccactgctt cagaatatgc ctacatgttt   420
gaaccagtaa aaaatgctga aaaaatttct aaaggggaac tacctaaaga agaattgggc   480
attaaagcaa tcaatgatca tgaattagaa atcactctag aaacagcaac accatatttt   540
gacgatttat tggctttccc ttctttcttg ccgcaacgtc aagatatcgt tgaacgtttt   600
```

```
ggtaaagatt atacaaagag tagcgataaa gcagtctaca atggtccctt tacgctaact    660
gagtttgatg gtcccggaac agatactaaa tggtctctaa ctaaaaatga agaatattgg    720
gacaaagaga cggtcaagtt agataaagtc gctatcaacg tggtgaaaga agccccaact    780
gccttaaatc tttatgaaac tggtgaagtg gatgatacgt atttatctgg cgaactggct    840
caacaaatgc aaaactcgcc tgacttggtc caattacaag ccgcttcttc tttctattta    900
gaaatgaatc aagcagatga aaaatcacca ttgactaatg caaacttacg tcgtgctatg    960
tcttatgcca tcgatcgcga ttcattagct aaaaatattt tagctaatgg ctctcttcct   1020
tcacaaggct tcgttcctgt ggatgtcgcg aaatcaccaa aaacgggtga agactttgtt   1080
aaagaagccg gcagcgacaa attagtcaaa tacgacaaga aaaaagctgt ggaatactgg   1140
aacaaagcga aacaagaact tggtgtttcc aacttaacg ttgatttaat ggtagacgac     1200
tctgaaggcg ctaaaaaat gggcgaatat cttcaaggat cactatctga tactttggaa     1260
ggcttaaaag taactgtgac gcctgtccct atggctgttc gcttagatcg tacctaaaa     1320
ggggatttcc aaatcgctgt tcgtggttgg agtgccgact attcagaccc aattaacttc   1380
ttagatttat tagaaagctc aacttctaat aaccgtggac gttacagcaa tcctgaatac   1440
gataaattca ttgctgcgtc caaaaccaca gatgttaacg atcctgaaaa acgctgggaa   1500
gatctaatca acgctgaaaa aacagtaatt gctgatatgg gtgttgtgcc aatttaccaa   1560
aaagcagaat cacacttacg tgcaccaaac gtgaaagaaa ttatttatca tccaacaggt   1620
gctaaatacg acttcaagtg ggcgtataaa gaa                                1653
```

Seq ID 77
```
atgtcaaaaa aaccatcagg ttaccaaaaa aaagttctca cggtgttgtt ggtatcgttt     60
ggctttggcg ggttgttagc gcagttttcg ccgatgaaat caggtgggtt ttcaagttgg    120
tttgatgatt catatcgtca agctagtgct gaatcaagta aaacaaaaga accagcacca    180
gttaaaattg agaaaaaagt caaaacctta agctatgtca acaggtcaa tcaagaaatt     240
gaaaagaagc aatatgatgg acatctggat ttgccgttag aattgcagac agatgctaaa    300
tggaaagaca ccgcatacgg atttggcaat gtggataagc cgaatacaat cgaaattaat    360
ggctgtgcga ttgtatcgct tgcaatggtt ggttcataca tggatcacca agaagttacc    420
cctctggatg tgttagcttg ggcaaaaaat gacttcttta tggaagggca agggacggcg    480
tggtctattt ttagtgcata tgctgaaatg aaaggctata actgtcaaga aattggggat    540
attgaaacag tggcagcttt cttgaaggaa ggtcatccag tcattattc tgtaaaacg      600
ggctatttca ctacaactg tcacattatg gtgatgagtg gtgtggatga aaaaggcgat    660
ttctggatta atgatccaaa cgattcagaa gaaaagggcc attcaaaacg gacatttaca   720
gccgaagaag tgatgaatga agcgttaaac ttctgggcat tttat                   765
```

Seq ID 78
```
ttgaagaaaa gtgtattatc ggctctaatg gtatgttcca ttacattaac aagcgtagcg     60
ttgccatccg cagcatttgc agatgaatac gatacaaaga ttcaacaaca agatcaaaaa    120
attaatgcgt taactagcca aatgtcagat gcagaagcaa aagttgccgc gattgaaaat    180
gatatggttg aaacggccaa acaaatcgat acattaacag ctaaaaagaa caagctatca    240
tcagaagtat ctaaattata tagtgaaatt tctgatttga atgtccgtat tcaaaaacgt    300
gaagtacaaa tgacaaaaca agcacgcgat gtccaagtga atggtcaaag tgattcaatt    360
attgatgctg tcttgatgc agattcagta gcagatgca ttggtcgcgt tcaagcggtc      420
tcaacaatga tgagcgccaa taatgaatta ctagaacaac aaaaagaaga caaagcgact    480
gttgaaaaga aaacaaagaa tgttgaaaaa caaattgctg aattagaagc agcaacaaaa    540
gaattaaatg ataaaacaga atcattaaaa acattgaaga ttcaacaaga agtggctaaa    600
aatgatttag aagcacaacg ttctgaagaa caagggaaaa aagacggctt cattaaacag    660
aaaaaagaag cggaaaaacg tttagcagaa gaacaagcac gtcaacgtgc agctgctaaa    720
aaagcagaag aacaaacggc agcgcaagca caagccgcaa cacaaaagcg ggcagcagag    780
caagcgaaag caacaaaagc agccaatgag gcagcagcat ccgccgctga agaaaaagcc    840
gcaacaccag tagtagaatc atcaacaact actgaaagta ccacaacgca agaaacaaca    900
acttcaagta ctgaaacaga aagtgttgta acaacacctg tggcagcagc acctgaaaaa    960
gaaaaagaag ttcctgtaac taacccaacg actcctgaaa aaggcaatga agcaaaacca   1020
gggaatggtg gcgtgacttc aggaaaacaa gccgcaatta atgcagcttt agcagatgtt   1080
ggtaattctt atgcaacagg ttggaatcaa ccaggcgaat gtttagtatc tgttcgtcgc   1140
tggttagcgc caggcggtat taacttcggt tatggttggtc caaacagtgg ctatgttgct  1200
tctggcgcaa cacaagttag ctggagtaac gtgcaaccag gcgatgttgt tcaatatgaa   1260
agtgcctata gtccagtag ctggattggt ggcgtgcata ctgtcttggt tacaggtgta    1320
agtggtggaa gtgttcaaat cgttgaagcg aacaacccag gaggttctgg ttatgtttct   1380
tctaactcta actggtcacc agcaccacca gcaggattcc gcgctgttgt atggcgtttt   1440
ccaggc                                                             1446
```

Seq ID 79
```
atgtcctcat tattaaaacg attggttcag ttggttttgt tagtcgtcgc tgtcttgctg     60
attcgacact atgttttctc ccctgctgcg gtgaacggct cttcaatgga accaacactt    120
cataacaacg accgtttatg gggtgacctcg attaaaaaac cacagcgctt tgatatatc     180
gctttcccta gtcctcgcaa cggccaacga gtagccaaac gtttaattgg tttacctggc    240
gaaacagtcg agtatcgcga tgataccctt tatattaatg gtgtatcact cagtgaagat    300
tacttagcaa gtgctaaacg aaatgtctct aaaaatgaaa attataccca agatttttacg   360
ctagagacct tagaagccac ccaatccctg accgttccag aaggcatgta ttttgtcttg    420
ggggataatc gcccgcgctc agacgacagt cgttattttg gctttgttaa acaagcgagt    480
gtggaaggtg ttttgacttt tcgttattat ccattagata aaattggctt tcca          534
```

Seq ID 80
```
atggcgaatg atgacttaca acaacgtctg gataaaggaa tgtacggtac gcccctcgtt    60
aatcctgaag aacaacacaa atacatgggc acatttcgtg aacgctgccg tctctccatg   120
accgtcgctg aaatgaaaga tgcacagaac caaaaacatt tattagaaga attagcgaag   180
catcctgaag ttacagtact tttaaatggc gaaatctctt ctgacctgca aagtacctat   240
attaaattat taaatcagca tggcgcaaac tttaagattt ttaacaattt tgtggagaac   300
aaccctgact ctctcggctt actgttagct gaaaaacatg cagtcgatga acctgtgatt   360
gacgttacag aaaaatatcc tcaagcaaca gaaacgccaa aagaagagcc aacagctaaa   420
aaaagctttt ggcaaaaact atttcactca                                    450
```

Seq ID 81
```
gtgtatttaa aacgaattga aattacagga tttaaatcat ttgcagataa aaccattatt    60
gaatttgaag atgatgtgac agcagtggtt ggtccgaatg gaagtgggaa aagtaatatc   120
acggaagccg ttcgctgggt attaggtgaa caatctgcga aaaaccttcg cggtggcaaa   180
atgaacgaca ttatctttgc cggttcggaa ggtcgtaagc cgctaaatat tgcggaagtt   240
accgtgacgt tagacaacag cgaccattat ttagcattgg attatagtga gattagtgtc   300
acccgtcgtt taaaacgaac gggcgaaagt gacttttta ttaataaaca agcgtgccgt    360
ttaaaagaca ttcaagattt atttatggat tcagggttag gcaaagaatc tttttcaatt   420
atttcacaag ggaaagttga agcgatttt aatagtaaac ctgaagaccg tcgagggatt    480
tttgaagaag cagcgggcgt tttaaaatac aaacaacgaa agaaaaaggc tgaacaaaag   540
cttttttgaaa ccgaggacaa cttaagccgt gtccaagata tcatttacga gttggaagac   600
caattagttc ctctagccgc ccaagcagat gcggctaaaa agtatttagc tttgaaagaa   660
gaacttaccg aaattgatgt caatttaact gtgaccgaaa ttcaagaagc aaaagctatt   720
tgggaaacga aaacgcaaga attaacagcc attgaagaaa agttagctgg cgcaagtaag   780
caagtccatg acttagaagg caaattggta cgtttacgca gcaagcgcaa tcgtttagat   840
gaacaaattg aaacggaaca acaacaatta ttacaagtga cagaagcgtt gaaacaagca   900
gaaggacaaa aaaatgtttt aattgaacgt tccaaacaca cctcccaaac ggcaagcgaa   960
tacgaagaga cattggcaga aacggccgaa aaaattgttc gttatcgaga agaattgcag  1020
acgctggaaa cagcaatcgc agaaaaaacg gcacaacgtc agacgttaaa agaagcattg  1080
gctttggcaa cgaaagatgt ggaaaaatat agtaagtctc ctaaagaatt aatggaagaa  1140
ttgcgcagtc agtatgtgga agtcatgcaa gaacaagcca atacagccaa tgacttaaaa  1200
tatttagaac gacaatatca acaagaaaca gctaaaaatc aacaatcgtt agcaaaacac  1260
gaagcgctag aagaacagat ggttgaagct ctcgcaatga aagagacgct agaaaaagag  1320
caaaaagttg caaaacaagg cttacaagaa cagttggaag aatacactgc gttgaaagca  1380
acgcttgaag ctaagcgtga acgattagca cagcgtcaaa acgatatgta ccaagcgtg   1440
aatcaagtgc aacaagcgaa ggcacgccaa aaaagtttgc aagaaattca agaaaattac  1500
gctggcttct atcaaggcgt gaaagctgta ttgcgtcaca aaaaccaatt aactgggatt  1560
gttggcgcag tggctgagtt aattgaagtg cctaaagaat atacgttggc gattgaaacg  1620
gccctaggtg gtgcggcgca acatattgtt gtggaaaatg agaaagatgg tcgagcaggc  1680
atcaccttct aaaaacaaca acacagtggc cgagcaacct ttttaccatt gactaccatt  1740
aagccacgct ctgtatcggc gatggttcag aatcgtttgg ctggtgcacc gggcttcgtg  1800
gggattgcca gtgaactagt tcgttatcca gaacaagtac aaaccgttat tcaaaatctc  1860
ttaggcgtca cgatttttagc ggcagattta acgagtgcca accagttagc aaaaattagtc  1920
aattaccaat atcgcgttgt ctcattggaa ggcgatgtga tgaatcctgg cggttcgatg  1980
actgggggag ccaataaacg tggcaaccaa ggaagcctat tttcacaagc gcaagaactt  2040
caaacaatta ctgaacaaat gactcaatta gaaacacaac tgagaagtgt agaacaagaa  2100
gtccaagcgc tctcgcaaga agtgaaaaca gccacggaac gtgcagaaat gttgcgctct  2160
gctggtgaac aaaatcgctt aaaacaacaa gaaattgaca ataaattagc caatcaaaca  2220
gaaacgatta ctcgtttaac aaaagaaaaa cgcttgtttg aatatgaatc acgggaattg  2280
catcaattct taaccgaata tcaaacgaaa aaagccacat tgacagagca acaagcaaac  2340
ttaacggcaa cgaaagaacg tttagatgct gaaatgaaac aagtggaaca agaagctagc  2400
caaatggaaa cctttaaagc ccaagcgcaa gaacgcttga cgacagttca agcagaacaa  2460
gcagtagcgg cggaacaatg tgcccatttt gctcgtcaaa agcaagacaa acaagaacaa  2520
ttggacgaat tgttgattcg agaaacggcc atccgtcaac agctacaaca attaagtagc  2580
cattctagcg atcaccagtt aaccgaagaa ggcttagcgg cacaagtcgc tcaattagcg  2640
gagaaacaaa cagcgttgca aacctcttta caaacggcac gtagtcaacg acaagcgttg  2700
caagaagaag tagacgagtt agatacaaaa ttagcggaag aaaacaaacg ccaacaacag  2760
tatttagcag ataaaacaca aatcgaagtc ctaaaaaatc gtgcagaaat gcaattagat  2820
agctcgttaa gttacttaca ggaagaatac agcttaacct tgaagctgc ctatgaagcg   2880
tatttcccaa tcgatgattt agcgcaagcg cagcaaacag tgaaacgaat aaaacaagaa   2940
attgagcggt taggacctgt taacttaagt gccatcgagc aatttgaaca agtcgatgag  3000
cgccatcaat tttttagtcag tcaacgagat gatttattaa atgcgaaaga acaattattt  3060
gaaaccatgg atgaaatgga ccaagaagtg aaagaacggt ttaaagaagt ctttgaagca  3120
attcgtggcc agtttaaagt ggtatttcct aatatgtttg tggcggacg ggcagaactg   3180
gtgctaacca acccagagga tctcttgaat acaggattg agattgaggc gcaaccacca   3240
gggaaaaaat tacaacattt aagcttgttg tctggaggcg aacgggcatt aacggcgatt  3300
gcgttactct tttcaattat tcgtgttcga ccagttccgt tctgtatttt agatgaagta  3360
gaagcagctt tagatgaagc caacgttgct cgttttggcc attatttaag cgaatttgaa  3420
gatggtacac agtttattgt ggtgacgcat cgtaaaggga caatggaagc cgccgatgtt  3480
ttatatggcg tgactatgca agaatctggg gtttcaaaaa tcgtttcagt tcgcttggaa  3540
```

```
gaggtcaaag aaggtggcgc aattgaaaaa agcaat                      3576


Seq ID 82
atgcaaaaag tagccaaagt tattgtggat gtaccaacga tgcaaaccga tcaaccattc   60
acttatttag ttcctgaaaa tttaaatgag caattggctg tcggcatgcg tgtagaagtg  120
ccttttggta acggtaatcg gcatgttcaa gggtttgttc tagccattga accagtggct  180
gcaacggttc ttgacgaaac caatgttcag ttgaaagaac tagtggcggt gttggattta  240
aaacctgttt taaatacaga aatgttggct ttagctgatt atatgaagga aaaaacgttt  300
gcttttaaaa tcacttgttt gcaaaccatg ttgcctagcg taatgcgtgc cgattaccaa  360
aaatatattt atttaaccgga tgagctttct gaagaattac aagatcagtt attttatggt  420
ttagaggaaa tttcgtggga ccaagcgcaa gaacgtgggc ttttgcctca actaatggcg  480
ttgcgtaaac agcaaaaagt cgatattcgc tacgaagtta cgacgagaaa taaagtcaaa  540
atggtccgct tcattcaagc cgctaaggaa tttgaacaac tagaggaaat tcgtttgggg  600
ttacgcaagg gagctaagaa gaaagagcaa cttctttact atttacaacg cttgggcact  660
gaaaaagtca ccgctgtgaa ggaaatgaaa gaattaggtt tcagtaccgc gcttttaaat  720
gaagcggcca aaaacggctg gttaacgttc attgaaaaag aagcgtatcg tgatccgttt  780
gctaatcaga cgtttgaaaa aacgactgcg ttatctttga atgcagaaca acaggtggct  840
gtggaaacaa tcttacaatc ggtccaagaa cagcaaagca aaacctactt attggaaggt  900
attacaggaa gcgggaaaac agaagtttat ctacaggtca ttgcagaagt tcttaaccaa  960
ggcaaaacag ccattatgct tgtccctgaa atttctttaa ctccgcaaat ggtgcaacgc 1020
tttaaaagtc gctttggcga acatgttgct gtaatgcaca gtggcttatc gcaaggtgaa 1080
aaatacgatg agtggcggaa gattgaacga ggagaagcgg aagtggtggt tggcgcgcgt 1140
tcagccattt ttgcgccaat cgaaaacatc ggcgtgatca ttgttgatga agaacacgag 1200
gcaagttata aacaagaaga aacccgcgc taccacgcga gagatttagc gatttggcgt 1260
tcggagtacc atcattgccc tgttgtttta ggcagtgcca cgccatcctt ggaatcacgc 1320
gctcgcgcac aaaaaaatgt ttatcagcgc ttacgattaa cacagcgtgc caatcaagcg 1380
gcaaccttgc ctacgattga tgtagtggat atgcggcaag aagtagaaaa tggcaatgtc 1440
tcttcttttt ccatgtcgtt gcaagaaaaa ctccaagaac gcttagagaa aaacgaacaa 1500
agtgttctcc tacttaatcg ccgtggctat tcgtcatttg tgatgtgtcg cgattgtggc 1560
tatgtgttgc cgtgtccgaa ttgtgatatt tccttaactt tgcatatgga tagtaaaaca 1620
atgaaatgtc attattgtgg acatgaagaa cggattcctt atcgctgtcc taactgtggc 1680
caagataaaa ttcgctacta cggaacaggc acacaaaaag tagaagaaga attacaaacg 1740
ttgttaccag acagtcgtat tctccgaatg gatgttgata cgacccgtcg caaaggtgca 1800
catgaaaaga ttttgcgaac atttggtgaa ggacaagcgg atatcttatt aggcacacaa 1860
atgattgcca aaggactgga ctttccaaat gtgacgttag tcggtgtctt aaatgctgat 1920
accgccttga atttacccga ttttcgttcc agtgaacgga ccttccaatt attgacacag 1980
gtcagcggtc gagctggtcg cgctgaaaaa ccggggcagg tcatcattca atcttttaac 2040
cccgagcatt atgcgataca attggcaaaa gcccaagatt atgaagattt ttatacaaaa 2100
gaaatgtata ttcgtcatcg tggggattat ccgccgtact attttacggt ccaaattacg 2160
gctagccatc ctgaagaaaa tgaagcagcg aaacaaatgt ttcaaatcgc gacgaaatta 2220
aaacaaggct tatcgccgca agctattttg ttgggaccaa cacccaacgc aattatgcgt 2280
gtcaataatc gatactttta tcaagtgatt atcaagtata aacaagaacc aatgcttcaa 2340
ccattattga aagaaatttt aaccgataca caacgagcaa cggctcgtgg cttaaagctt 2400
tccattgatg cagaacccat gaatttttatt                                 2430


Seq ID 83
atggacttac atttaaccgt tgatctaaaa aaaacaaaga cactcattga cttagccgat   60
caggcccact taatcattga acaaacaact aaattaccga aaaaagaggc agcgattctt  120
ttcttgctaa tagataacca actagttgct ttaggacttg ctgaagaaaa agctacatat  180
aaagaagttt cttttgatca ctcgatttta ttgaaaccga cttgggacac agaacttgcc  240
tatttagcgc aagctttcct agatgatgcg aaagaaagtg gtttaacatt ggaagcaacc  300
ccaacgacag tcaaaattcc gaaaagtgcg gtagcgacag tgactaacta taaagaaaca  360
gtcaccttta tcttggaacg ttttggctac tctctgttta aaaaaccggc gccgaagaaa  420
gccagaccag ccaaagcacg ccacaaatgg acaaaagaag ttagccaaat tcctttttat  480
attgatactc gacaaagcaa agcaactgtt ttttggcaaa agcgcaatga aatgcttaat  540
aaagctggcg cgacaatgat gcccgaagcc cctttaaata aagacgggtc tgtcggtttc  600
tccgctagat ttggtgaaaa acttcgtgat gaacgcaagg gccaatttaa agactttgtc  660
acaacggaag atatcgtctt aaaaagcgtt aatgaagtcg gcttgtttct ctattttgct  720
ggaacaaaca gctggttaga gctagttgat gagaacggca aaaccttaaa tgaatggact  780
gtggtggaa                                                          789


Seq ID 84
atgaacagat ggaaagtata tgcaacggta atcgcttgta tgttatttgg ctggattggc   60
gtggaggcgc acgcttctga atttaatttt gcggtcacac caacaattcc cgaaaatcaa  120
gtggataaat caaaaaccta ctttgactta aaaatggcgc ctggtgccaa acaaaccgta  180
gaaattcagt tacgcaatga tacagatgaa gacattacca ttgaaaatac ggtgaactca  240
gcgacaacaa atttaaatgg cgtagtagaa tatggccaaa acgggatcaa acctgacaaa  300
accttacgtt ttaacttaaa agattatgtg gaagcaccga aagaaatcat cttgccgaag  360
cattcccaaa agaccttacc tttaaccatt acgatgccta aagattcttt tgatggcgtg  420
atggctggcg gtataacact caaagagaaa aagaaagaa caacgacttc tgcggatcaa  480
tcaaaagggt tagctattaa taatgaatac tcctatgttg tggctattat tcttcagcaa  540
```

```
aatgagacaa aggttcaacc agatttaaaa ttactggggg ttaaaccagg ccaagtcaac   600
gcgcgaaacg tcatcaatgt ttctttacaa aacccacaag cggcctattt aaaccaatta   660
catttaatca acactgtttc aaaaggaggc gaaacgcttt accaatccga tactgaggat   720
atgcaagtgc cgccaaactc taactttagt tacccaattt ctttaaaagg ggaacgatta   780
acgccaggaa aatatgtctt gaaatcaacg gcctatgtga taaaagatga aaagggcacc   840
tatcaagtca aaggcgccaa tggtgaagaa cggtacctgt acaaatggga atttacaaaa   900
gaatttacta tttctgggga cgtcgctaaa gaattaaatg aaaaagacgt aaccattaaa   960
ggaaccaatt ggtggttgta tctactgatt gcattaatca ttctagcgct gctcttattg  1020
attttcttct tgtatcgtaa aaagaaaaaa gaggaagaac aacaatctga gcaa         1074

Seq ID 85
atgagcaaga attttgggc aacattgcca aagcccttt ttgttttagc accgatggaa    60
gatgtcactg atgtggtctt tcgtcatgtc gtgaaagaag ctggggcacc agatgtgttt   120
tttacagaat ttactaattc ggatagttat tgtcatcctg aaggaaaaga tagtgtacgc   180
gggcgcctcg tttttacaga agatgaacag ccgatggtgg cacatatttg ggggggataaa   240
ccggaatttt tccgtgaaat gagtatcggc gtagcggaga tgggctttca agggttagac   300
ataaatatgg gctgtcctgt gcctaacgtg gctgagcgtg gaaaaggcag cgggctaatt   360
ttgcgcccag aagtcgcggc tgaattgatt gacgctgcca aagcaggtgg cttacctgtt   420
agtgtcaaaa cacgcattgg gtttactgaa atggcggaaa tggaggcgtg gatcacgcat   480
ttattagagc aagacattgc gaatctatcc attcatttgc gaacacgaaa agagatgagc   540
aaagtggatg cccattggga ggtcattccg caaattatgg ctattcggga ccgtgtcgca   600
ccgcaaacga caattacgat taatggggat attcccgatc gtcaaaaggg cctagaatta   660
gcagaacaat atggtgtaga tggcatcatg attggtcggg ggatttttaa aaatccttat   720
gcctttgaaa aagaacccaa aacacatacg ccacaagaat tgctgggctt gctacgttta   780
caattggatt tgcaagacaa atatgcggaa ttggtgcctc gctccatcgt tgggctgcat   840
cgcttcttta aaatttatgt caaaggcttt ccaggtgcca gtgattaag agcacaatta    900
atgaatacaa aatcaacgga tgaagtgcgc cagttgttag cgacgtttga aacagaacat   960
ggtgtgcttg ac                                                       972

Seq ID 86
atgcgttatt tagatgcaga agcaattgag aatatagcca caggagcagc ttttttaggg    60
acaggaggtg gcggtgatcc ctacattggt aaaatgatgg ctttgtccgc catcgaagaa   120
aacggacctg tcaaactggt ttctccagaa gaaattgccg cggaggattt tttcctaccc   180
gccgcaatga tgggtgcccc atctgtcgca attgaaaaat ttcccaaagg cgacgaattc   240
gtccgtgtct ttgagaagtt aggaaaatat ttagaccaag aaacgattgc gggaaccttt   300
ccaatgcaag ctgggtgcgt caattcaatg attccaattg ttgttgcagc gaagctaggc   360
attcccttgg tggattgcga tggtatgggt cgggcctttc cagaattgcc tatggtaacg   420
ttccatttga atgggatgtc agcgacccct atggcaatta ccgatgaaaa aggaaatatt   480
ggcattatgg aaacgattga taatacttgg acagaacgtc ttgctcgtgt tcaaacggtt   540
gaaatggggcg ccagtgcttt agtgagcatt tatcccgcga caggcaaaca attacaagac   600
tatgggattc acaacatcgt gacattatca gaagaaattg gcaaagtgat tcgaggtacc   660
tatgcagatg aacaagaaaa acgccaagca ttagtagaag ttacggatgg ctttgaattg   720
ttccaaggaa aaattctaga tgtgaacga gaagtaaaag gtggcttcaa tttgggacgt   780
gtcaaattga gtggcttaaa cagtgacgct ggttcagaag cagtcgtcca ttttcaaaat   840
gaaaatttaa ttgccgagaa agatggtcag gtgattgcga tgacgcctga tttgatttgt   900
atggtagatt tagaaacttt aacgcctgtg acaacagaaa gcttgaaata cggcaaacgt   960
gtccaagtaa tgggcttgaa agcgaatgcc gcttggcgaa cgaaaaaagg tatcgaaaca  1020
gtgggtcctc ggtatttcgg ctacgaaatg gattatcaac cactagaaaa cttagtagca  1080
aaggaggaca aa                                                      1092

Seq ID 87
gtgacaattg ttttattact actattaaat gtctcaatta tgttccttt aggacaacta    60
cttagtgaaa actcgccctt agctggtgtc atctgggcca ctaaggatgc ggtcgctaaa   120
gaattaaata ctgatttgtt ttggagctgg aataaatttg tcttacccat tttctttta    180
gtagatgcag ctgttctcta ttggcgctta attcgtcgct accatcaaat gcaattgcgt   240
catattatca gtgagttaca ttacatcgct gatggtaatt acaatcaccg cattccgttt   300
gaattaagtg gtgatttagc aaaagtcgtg acgagtatta acggtctggt ggatagtact   360
gtcgcagcca ttgaagatga acggcgaatt gaaaaatcaa aagatgaatt aattacgaat   420
gtcagtcatg atattcggac accgttaact tctattattg gttatctggg attgattgaa   480
gatcgtcaat ttcatagcca agaagattta ctgaaatata cgcataccgc ttacgtcaaa   540
gcgaaacaaa tgaaattatt agttgatgat cttttgaat atacaaaagt ccgtcaacca   600
agtgtaccga ttcatacgac gacttttgat atggcacaat taattgaaca gcttgctgcc   660
gattttgaat tggaagcgaa aaagattaat atgcaaatcc aagtaaaagc caatccagct   720
tctttaatga tggaaggcga tacggaaaaa ctagtccgag tctttaacaa ccttcttca    780
aacgctttga aatatggcaa aggtggccat catattgtaa tggaagtaga caaagtcggc   840
accgaagcaa ttattgctgt tcgcaatgat ggaccagcta tccctaaaca ttctttagat   900
cagttgtttg accgttctta tcgtgtcgaa gaatcccgat cacaagaaac ggggggtact   960
ggtttaggtc tggctattgg acaaagtatt gtcgctttac atggtggata tatttatgcg  1020
aagtctgatc aaaagtggac ttcctttatt attcatctac ccttacagcg taccaacaag  1080
aaatcagaga gt                                                      1092
```

```
Seq ID 88
atggaagcag aagcattacg agcaatcgtg gctgagaatc gtcaattaga acagaatttg      60
actaaaagaa atgaacaata catttttgac ttaaaaaaat cattaaaagc ggctaattta     120
tctgaagaag agctagcatt ggctctacac ggtattttgc cagaattagt agccggtcaa     180
aaaacaggta aaacagcccg tcagttattt ggcactgttt cagaacggac ggaagcaatt     240
ttaaataaac cagcagaagt caaagaacca gcgggctgga tgatttggct cgacaatacc     300
ttactattac ttggtttatt aacaatcatg ttagcggcga tgtctctatt ttcaaaagga     360
accgcacaac ctctcggatt aaccacatat atcttaggtg cgatggctgg gggttatgtt     420
ttctatctaa tgcacaaata tgtttaccgt tttgatcgtc aaggtggcga caaaagcaaa     480
cgtcctggtt ggttaaaaac gaccgccatt ctatttggtg ggatgttctt gtggattgca     540
gttttcgctg gttcagcgat gctaccacca gtgatcaatc ccatcttaga tccgatgatt     600
gctttagtaa tcggtgcgtt agcctttgta gcacgctatt tcttcaagaa aaaatacaat     660
attcaaggta gtttcatgac acgtcaa                                         687

Seq ID 89
atgaatttag ttttagccga gattgccgcc atggggctga aaaacatttc gattgcacct      60
agttcaattg caaatgttca tgaacctctg attgaacaca ttaaaaatgg agtagtcacc     120
aacattactt cttctggttt gagagataaa gtaggagccg ctatttcagc aggtatcatg     180
gaaaatcctg tggtgattcg ttcacatggt ggccgggcac gggcagtagc agcaggcgat     240
atccatattg atgtagcttt cttaggtgca ccaagttcag acgcttatgg aaatgcgaat     300
ggaaccaaag gaaaagctac ctgtggttct cttggatatg cgatggtcga tgcaaaatat     360
gcggatcaag tggtcatcat tactgacaca ttagtaccat atcctaatac ccctatcagt     420
attccacaaa cagatgttga ttatgtggta gaaattgatg caattggtga tccagatgga     480
attgccaaag gcgcgacacg ttttacaaag aatcctaaat aattattgat tgctgaatat     540
gccgcaaaaa taattacaca ttctccttac tacaaagaag gcttctcatt ccaaacagga     600
acaggtggct cttcactagc agttacgcgg ttcatgagag aacaaatgct taaagatggc     660
attaaagcaa gtttcgcttt aggcggaatt acgaatgcaa tggtggaact tttggaagaa     720
ggcttagtag aaaaaattat tgacgtccaa gactttgatc atccatcagc tatttcttta     780
ggggaaaatg ccaatcatta tgaaattgat gccaatatgt atgcctcacc gcttagcaaa     840
ggggcggtta ttaatcaatt agacacagct attttatctg ctttagaagc ggacacagat     900
tttaatgtaa atgtcattac tggttcagat ggtgtcatcc gtggtgcttc tggcggacat     960
tccgatacaa gtatggcttg taaaatgagt ctagtgattg caccacttgt tcgtggacgg    1020
attccaacaa tcgttgagca agtaaataca gtcgttacac caggtacgag tgtcgatgtc   1080
gtagtaacag aagtcggcat tgcaatcaat ccaaaaagaa cagacttaat tgattgtttt    1140
aaaaccttag atgttcctca gtttacgtta gaagagttga agacaaagc ctacaacatt     1200
gttggcaccc cggaaccaat taaatatgga gacaaagtcg ttgcgttaat tgaataccga    1260
gatggtagtc taattgatgt ggtacgcaat gtt                                 1293

Seq ID 90
atgacgatga atcctttaa aggaaagcaa tttcagcagg atgtgattat tgtagccgtg      60
ggctactatc ttcgttataa ccttagctat cgtgaagttc aagaaatctt atatgatcgt     120
ggcattaacg tttctcatac gacgatttat cgttgggtgc aagaatatgg caaactactc     180
tatcaaattt ggaaaaagaa aaataaacaat tccttttatt catggaaaat ggatgaaacg     240
tacatcaaaa ttaaaggaaa atggcattat ttgtatcgag ccatcgatgc agatggttta     300
accttggata tttggttacg taaaaaacgg gacacacaag cagcctatgc ttttcttaag     360
cggttagtga agcagtttga tgaaccgaag gttgtagtca cagataaagc cccctctatt     420
acaagtgcct ttaagaaact aaaaagaatac ggctttttatc aagggacaga acatcgtacc     480
attaaatacc tgaataattt gattgaacaa gaccatcgtc cagtaaagag acgcaataaa     540
ttctatcgaa gtttacgcac tgcctctacc acgattaaag gcatggaatg cattcgagga     600
ttatataaga aaacccggaa agaaggcact ctcttcgggt tttcggtctg tactgaaatc     660
aaggtattat tgggaatccc agct                                           684

Seq ID 91
atgaataaaa aaagaattat cttttgggctt ctctctttttt tcttgcctat attttttagtc      60
tttggaggct tactttttttt tctttttatta ttaacgagta cgtcagatac ttcaaaaaat     120
gattgtattc agccaagtat aaataatcca actgatgcga cagatacacc taaatcgatc     180
gagcagtttg taaaaagcca taaagatgct tacctttttat catggaaagc aggtggcttt     240
ttaccgtctg ctagtatttc tcaaacgatg gtagaaaatg ggtttaattt tactaatcca     300
tcggggacgt catttggca ggcgcacaat atgggcggtg ttaaaacgtc aaaaaaagaa     360
gattttcctg taactttagc aacattcggc caagattctg ttgatatttc tggtacaaag     420
ccagggtcaa acgtcggtga tggcactggt ggggcatata cctggtttaa agactacaat     480
gctggaattg ttggaaaagc agaatttatg gcacaccaga cactgtatac aggtgctatc     540
aataatactg acggattaag tactttatca gctatttatt caggaggatg ggctacagac     600
cctacttacc tcatgaagtt acaggccaca tataatagct taggcaagca gtttcaatgg     660
ttggaccaag aagcaatac agaaatatggt aacgcgccgt ttaaaaagag cgaacttgtc     720
ctaaatattc ccggaaaatc cccaatcaca aacgaaaat atggtaaaaa ttctgattgc     780
gttgttactt cagatacatc cgatcaagtt acaggacaaa atactgcccc atcattagaa     840
gtaccaagtg cgtataaagg aaagttaaca cttccgcta tcgattcgaa tgattacgct     900
ggaaataact atccttttgg acaatgtacg tggtacgcct ataatcgtat ggcgcaaatt     960
ggcaaaccaa ttgaatggtt ttcaggtgat ggcggaaatg gcgcaggttg ggctaattct    1020
gctcgtgcga aagggtacac agtagtaaaa gtaaaaccaa gtgttggttg ggcagctagt    1080
```

```
atgcaaggtg gcataggagg ttctgcaccg ccctatgggc atgtggcagt tgtagagtat  1140
gtcaattctg acggcagtat tcttgtgagt gaagcaaatg ttattaatca aggctctggt  1200
acgcgttcat ggcgagtatt agacagagca acggttgaac aaattgattt tattcaagga  1260
aagggagcg                                                          1269


Seq ID 92
ttgaacttac aaaattaaaat taaaatggag gaaataatga tgaaaaaaaa tcttttttgct   60
agtttattta gtgccacact actatttggg ggaagtgaaa tttctgcttt tgcacaagaa  120
attatccctg atgatactac gacaccgccc attgaagtac caacagaacc aagtacacca  180
gaaaagccaa cagatccaac accgccaatt gagccacctg tagaccctgt agagccacct  240
attacaccaa cggagccaac agaaccgaca gagccgacaa caccaacaga acctgaaaaa  300
ccagtagttc caacggaacc gacaacacca acagaaccga caactcctac agagccaagt  360
gaaccagaac aaccaacgga gccaagtaaa ccagtagaac ctgaaaaacc agttacacca  420
agcaaaccag cagaaccga aaaactgtg acaccaacta aaccaacaga acctgaaaaa  480
ccagtacaac cagcagaacc aagcaagcca atcgacgttg ttgtaacgcc aacaggggaa  540
ttaaatcacg ctggaaatgg tacacaacag ccaacagtcc ctattgaaac aagtaatttg  600
gcagaaatca cacacgtgcc aagtgtaaca acacctatta caactacaga cggagaaaac  660
attgtagctg tagaaaaagg tgttccactt acacaaacag cagaagggtt aaaacctatt  720
caatcgagtt acaaagtatt gcctagcgga aatgtagaag taaaaggtaa ggacggtaaa  780
atgaaggttt taccatacag aggtgaagaa atgaatatct ttttatctgc cgtaggcggt  840
atcttgtctg tagtatctgg gtttgtcatc tttaaaaaac gcaaagctaa agta         894


Seq ID 93
atgaagcaac aaacagaagt aaagaaacgt tttaaaatgt ataaggcaaa gaagcattgg   60
gtggtagccc ctattctttt tataggtgtg ttaggagttg taggattagc tactgatgat  120
gtacaagctg cggaattaga tacgcaacca ggaacaacga cggtgcaacc cgataacccc  180
gatccgcagg taggtagtac aacacctaag acagcagtaa ctgaagaagc aacagtacaa  240
aaagacacta cttctcaacc gaccaaagta gaagaagtag cgtctgaaaa aaatggagct  300
gaacagagtt cagctactcc aaatgatacc acaaacgcgc aacaaccaac agtaggagca  360
gaaaaatcag cacaagaaca accagtagta agccctgaaa caaccaatga acctctaggg  420
cagccaacag aagttgcacc agctgaaaat gaagcgaata aatcaacgtc cattcctaaa  480
gaatttgaaa caccagacgt tgacaaagca gttgatgaag caaaaaaaga tccaaacatt  540
accgttgttg aaaaaccagc agaagactta ggcaacgttt cttctaaaga tttagctgca  600
aaagaaaaag aagtagacca actacaaaaa gaacaagcga aaaagattgc gcaacaagca  660
gctgaattaa aagccaaaaa tgaaaaaatt gccaaagaaa atgcagaaat tgcggcaaaa  720
aacaaagcgg aaaaagaacg ctacgagaca gaagtcgcgg aataacaaca acataaaaat  780
gaaaatggct atgtagcaaa accagtaaat aaaacgctaa ttttcgatcg tgaagcaaca  840
aaaaattcca aagttgtttc tgtaaaagct gcagaatata tagacgctaa aaaactaact  900
gataaacata aagataaaaa attacttatc agtatgctta gtgtagattc aagcgggtta  960
acaactaaag actcgaaaaa agcacatttt tattataata acggtgcagg aggaacattg 1020
gttgttcttc acaaaaatca accagtaact attacctatg gcaatttgaa tgctagttat 1080
ttgggtaaaa aaattgctag tgctgaattc caatatacag tgaaggccac acctgattca 1140
aaaggtcgat tgaatgcttt cttacatgat gatccagtgg ccacaattgt ctatggaatt 1200
aacattgacc ctcgtacaaa gaaggctggt gctgagattg aaatgctcgt tcgcttcttt 1260
ggagaagatg gcaaagaaat cttgccaacg aaagagaatc catttgtatt ttcaggtgct 1320
tcattaaatt cacgtggtga aaacattacg tatgagttcg taaaagtagg aaacacggat 1380
actgttcatg aaaattaatg atcaaaagta gctcgtcatg gaaataaagt ttattctaaa 1440
acggatattg atgtagggac gaatgggatt tcaataagtg actgggaagc agttcaaggc 1500
aaagaatata ttggcgcaac tgttatttca acaccaaata gaattaaatt cactttcggg 1560
aatgaaattg ttaacaatcc agggtatgac ggaaattcga tgtggttcgc atttaatacg 1620
gatttaaaag caaatcaat tacgccttat caagaaaaag gacgtccaaa gcaaccagaa 1680
aaagcaacga ttgaattcaa tcgatacaaa gccaatgtgg ttcctgttct tgttccgaat 1740
aaagaagtca ctgatggcca gaaaaatatc aatgatttaa atgtgaaacg tggcgattct 1800
ttacaataca ttgtgacagg ggatacgaca gaacttgcca aagtagatcc aaaaacagta 1860
acaaaacaag ggattcgaga taccttgat gcagaaaaag tgacgattga tttatccaaa 1920
gtgaaagttt atcaagcaga cgcaagtcta aacgaaaaag acttaaaagc tgttgctgca 1980
gcaattaatt caggaaaaagc taaagacgtg accgcttctt atgaccttca tttagaccaa 2040
aacaccgtta cagcaatgat gaaaaccaac gcagacgact ctgttgtttt agcaatgggg 2100
tataaatatt tacttgtctt gccatttgta gtgaaaaatg tagaaggcga ttttgaaaat 2160
acagctgttc aattaacaaa cgatgggga acggtaacaa atacagtgat taaccatgtg 2220
ccaggtagta atccttccaa agatgtaaaa gcagataaaa acggtacagt tggcagtgtt 2280
tctctacatg ataaagatat tccgttacaa acaaaaattt attatgaagt gaaatcttcc 2340
gaacgtccag ccaactatgg cggaatcaca gaagaatggg gcatgaatga tgtcttggac 2400
acgacccatg atcgtttcac aggaaaatgg cacgctatta cgaactatga ccttaaagta 2460
ggggataaaa cgttaaaagc aggaacagat atttctgcct acattctttt agaaaacaaa 2520
gacaataaag acttgacgtt tacaatgaat caagcattat tggcagcgtt aaatgaagga 2580
agcaataaag taggcaaaca agcttggtct gtgtatctgg aagtcgaacg gatcaaaaca 2640
ggtgacgtag aaaacacgca aacagaaaac tacaacaaag agcttgttcg ttctaatacg 2700
gtggtgacac atacgcctga tgatccaaaa ccaaccaaag ccgttcataa caaaaaaggg 2760
gaagacatta tcatggaaa agtggctcgt ggtgatgttc tttcttatga aatgacttgg 2820
gacttaaaag ggtacgataa ggactttgcc tttgatacag tcgatcttgc gacaggcgtt 2880
```

```
tctttcttcg atgattacga tgaaacgaag gtgacaccaa tcaaagactt acttcgtgtc    2940
aaagattcta aaggggaaga cattacgaac cagttcacga tctcttggga tgatgccaaa    3000
ggcacggtga cgatttctgc caaagaccca caagccttta ttttggcgca tggtgggcaa    3060
gaattacgtg taactttacc aacaaaagtt aaagccaatg tttctggtgt tgtgtataat    3120
ttagcggaac aaaatacatt tggtcaacga attaaaacca ataccgttgt caaccatatt    3180
ccaaaagtga accctaaaaa agacgtggtt attaaagtcg gtgataaaca aagtcaaaat    3240
ggtgccacaa tcaaattagg ggagaaattc ttctatgaat ttacaagtag tgacattcct    3300
gcagaatacg ctggtattgt ggaagaatgg tcgattagcg ataaactaga cgtcaaacat    3360
gacaaattta gtggccaatg gtctgtgttt gccaattcta cgtttgtctt agcagacgga    3420
accaaagtga ataaagggga cgacatttcg aaactattca cgatgacctt tgaacaaggg    3480
gtagtgaaaa tcacagccag tcaagccttt ttagatgcga tgaatctaaa agaaaacaaa    3540
aacgttgcac actcatggaa agcgttcatt ggtgtagaac gaattgcggc aggagacgtt    3600
tacaacacaa tcgaagaatc tttcaacaat gagaagatta aaactaatac ggtagtgacg    3660
catacgccag aaaaaccaca aacgccacca gaaaaaacag tgattgtacc accaacacca    3720
aaaacaccgc aagcaccagt agagccatta gtggtagaaa aggcaagtgt ggtgccagaa    3780
ttgccgcaaa caggcgaaaa acaaaatgtc ttattaacgg tagctggtag tttagccgca    3840
atgcttggct tagcaggctt aggctttaaa cgtagaaaag aaacaaaa              3888
```

Seq ID 94
```
atgaaatatg aacgtccatt aaaaagagag tcacaaatca aggagtttga gttaggcacg     60
cacgcggcag taattgaaaa agttcaaaag aaacggtctc aaaaagggaa tgatatgttt    120
ttactctctc ttttgggaaa aagtaacgaa aaaggggttt attttcttac cttcgggaat    180
gattatactg aagataattt acgctatatt ttagccagta tccaagacaa cggtgtgagg    240
attccagacg ttgactttgg ctataatcga gaaacttttg aatttctgaa aggtaaagat    300
gtctatattc aggttgaaga acaagagtat aaaggaaagg ttaaacatgc ggtaacaaat    360
tttttaactc aggatgaatt tgaagaaagc gaagaaatgg agtttttcaga aagcaatact    420
gaagaagatt gg                                                       432
```

Seq ID 95
```
ttggtggaac aaaagagaaat tgtacaccgc tattatgttt taggtactag tactttagat     60
aaagtattaa aagtgctgct taatttatcc agtgatggtg tattggcgct aaaaagacaag    120
tttcttcctt taaaagggga aaataatttg tataagctaa tgaatcgtga ggatccactg    180
acttcggcgc aacttcatga aaaagtgaat ttaaataagt taaaagaaca gttggaagcc    240
caagggcttc cttttgcgtt taaagaaacc aaagaaggga ctaacttcta ttttcgagta    300
aaagatacgg aactcgcaaa gaaagcctta gaacgcgtat taacggatat taagaagaac    360
ccacaaatga ttttgcgtaa gccaaatacc atgacttttg atgaacggtt agcctacact    420
gcagccaata aaaaatacgt gggtaaaatc gatcaaacaa aagcaatcac aaaaggaagg    480
aaactg                                                              486
```

Seq ID 96
```
gtgcgtatga ataaaaaaac accaaaacca ttaaccccac ccaaagcctt ttttaactgg     60
tgtactgcac aaatacccac gtatgaatgg caaaacaaaa aggagacaat tcttgcgagt    120
tctcgcaaga attgccctat catacaaaaa cgattaacaa aatactcacg attatcgttt    180
ccaacgaaat tttactcctt tggaatcatt ctcgttcgtg caaaacgaat tgaaattcaa    240
acccatagtt attggcgacac gattacagat ggaatagaaa accttatcta tgagccgtca    300
aaccttgaac ggttttcaaa cgatacccat gttaaagccc attatgagaa cggacattgg    360
catgagggct tgttagcaaa ttatggtttt atgagtagtg cttatacaaa tacagtattc    420
tatccaaata actggcaaga aaaactaaac actgtatcgg aattaaagta cttacagtta    480
cctgaaatag aacgtcaaga actagctcat atttacaaat accgaaacga aattgagttt    540
ttacaaaaga taggtgccac aaccttagcg aatgaaatca tttttgacga ctatcgcaac    600
gttttttggac tgtatttttca taaagtagat atgagagtaa ttactaagaa atggctaaaa    660
gcgaataaac aaaaaattaa aacaagaaat ccaacctttc atgaattcat gttagaaaaa    720
acgttaaaag agcgaaatgc gcctatgata aacgaaatag aaaagtacgt gcattattcg    780
caaattaagc aactaccgaa agaagtgaac ttaacaaagt ttcaaaaatg gtttttatacga    840
aaaggggagc gttttgacta ttatatggac tatttacaca tgttagaaga gttgaacacc    900
ccacttaata acgatagtgt attgtatcct gaaaacttac aagtagctca cgataacgct    960
atgaacacgc taaacctttt gaaaagtgag attgaagaaa agcagtatca agaacggaaa    1020
aaccaaatta aagcgttaga agctgaaatt gacgatttgc ttttctcac accacattca    1080
ttacaagaaa tcattcaaga aggaagtatt ctacgccatt gtgtgggcag tcaacactat    1140
atcgaacggc atacgcaagg aaaaacaacg attgttttta tccgaagaaa agaaaagcca    1200
gatatgcctt acttcacatt agaatatcga aatcaacagg ttatacaaat acaaggaaaa    1260
tgtaaccgtc aggaagtacc agaaaagata aaacaagcag tcagacaatg gcaagagaat    1320
ttacagcacg ccttatcatc a                                             1341
```

Seq ID 97
```
ccaagtgaaa agggaaacca aggaagggtt gaaaagagaa gcgacgcaaa aagtacgaac     60
cacgttaacg caggaagata cactgaa                                        87
```

Seq ID 98
```
gagaacgaat tcaagctgca caaaattggt ttacattctt ttctcgatcg ggtacaccaa     60
tcggtggttc tgggagtgag gtctttgcgc agtataaaga aaaaatgcaa ccattaccca    120
```

```
cagacagaga  aacaaaagca  ggacaaggct  ggccaggcaa  tgcgtatgcg  ccagggaatt     180
gtacttggta  tgtgtttaat  cgttttgcac  agt                                    213


Seq ID 99
aacatttttt  ccaaaagcag  tgctcctatt  attcggcgct  tcattaggtc  aatggctagg      60
tgggatgttt  gctttcttga  tgataggtcg  cttgatgcag  gtgccagttt  cgatggtttc     120
tgtttaccca  atgtttgtta  tcgctacgtt  gattgggatg  ttaactatgg  ttcctggtgg     180
tatgggaacg  tt                                                            192


Seq ID 100
aggggaagat  gttacatatt  taacgaagtt  cccaggtgtt  ggcaagaaaa  cagcccaaca      60
aatgatttta  gatttaaaag  gaaaattagg  cgaattagag  tcttccgaag  cagcggttgc     120
ggccatgacc  gcaaccgaag  cggtgacaac  aagcaatcaa  gcattggcag  aagcgttaga     180
agcattaagt  gccctaggct  acag                                              204


Seq ID 101
accagattgt  ggggaagaca  gctatgtcgg  aaatcacaaa  ttagaaaatc  gtcgggtatt      60
aattacaggc  gg                                                             72


Seq ID 102
cttggtttgt  catttataaa  acacgttttg  gcttacggtt  acgttctgtg  ggggaacatc      60
cacaagcggc  agatacattg  ggaattaatg  tctact                                 96


Seq ID 103
cacttggttt  tcgtccatgt  aaaagatgtc  aaccaaccgg  gcaagtcgtt  cctaacgagg      60
aatgggttgc  tcaaattaag  caatttattg  atcggcattc  agcggaacca  t              111


Seq ID 104
aagacggtag  tggcgcggca  gctgcgatgg  aaatggcgat  tgcagaagca  gggattacac      60
cggaacaaat  tgggtatatc  aatgctcatg  gtacgtcaac  cccagcgaat  gacgctgcgg     120
aaacaacggc  catcaaacgt  gtctttggtg  agcgagcatc  gcaaattcca  atttcaagta     180
caaagagta                                                                 189


Seq ID 105
tgctacagca  aatttaggca  aacatcaaaa  caatggccaa  acaagaggcg  acaaggcgac      60


Seq ID 106
gtggctagtc  cctcgaaaag  tgcgccattt  tttacgcacg  cgaaaaaatg  tcaaaatcaa      60
cggagaaaca  gctgcttttc  attttccagt  ccaagctggc  gattgtgtca  ccttaacgtt     120


Seq ID 107
aacaaggggc  ggcagccaat  ttacagaaag  ctccagcgcg  caccttttgat  cctaaagtat     60
ttaagcccag  cgcaagtaaa  ccagcttata  cacaaccagc  tacaaaatca  gttagcaaca     120
aagtggcaag  tggtggcgaa  aaagtttctt  ggcaagcagg  tgataaggtc  caacataaag     180
catggggcgt  tggtacagtc  gttcgtg                                            207


Seq ID 108
gagagcatca  aggatttcaa  caacgtttcc  agaaaaaaac  agataaaaag  agtccaatca      60
aaatggaaga  ttatctttta  ctcagtcgtc  atttcactgg  actcagccgt  gaaacctatc     120
gggaagttct  ccgatatttt  caaagaggcg  cttttcagca  atactactcc  gaagaccaag     180
aagtcatcag  tcaaagaata  cgtgaaaatt                                         210


Seq ID 109
cacctgtggt  tggtgctgcg  gtacctgccc  aagcagaatc  aagccagcaa  gcatttatta      60
atgaaatcgg  aaattcagca  gcagccgttg  ccaatagcaa  tgatatgtat  gcttccgtca     120


Seq ID 110
gtcatgtttc  ggcaaaatca  caggcgcgac  aagaagaaaa  agaggcgaag  cgagcagcga      60
agaaagaagc  caaagctgca  gaacgccaag  caaaaattga  agccgctgcc  caacaaaaat     120
tacaagagcg  ggaacgcatg  gaacaagcgg  cagccgagcg  ct                         162


Seq ID 111
aggattgcga  attggaattt  ggaatatacg  aataatttgg  aggtaatact  tttgggtaaa      60
aaaggttcaa  agattaaaaa  gaaaaaacgc  cgacttaaag  aaaaagctat  tgcaaatggc     120
acatacagca  agcgagggaa  aagcgatggt  atgtataaaa  tgcaagggac  agatgattgt     180
ttgggaaaaa  ga                                                             192


Seq ID 112
aaaacgaata  cggcattggc  gggcatggct  acattggcta  tgatgaatgg  cacgacggca      60
aaggcatcaa  gctgtcccgc  gccgacgatt  tttccggtgg  caattatgat  acggttacgc     120
```

Seq ID 113
cgattgcccg cgccgccgct gaggaaaacc cgcagtttgc catcaatgcc gaaaagctgg   60
aacaggtgca gcccaaagac c   81

Seq ID 114
ggaggtggaa cccatttgaa acaatatgaa atcattctgg ccgaccctcc gtggagatac   60
tcacaaaagg ggggtgcagg gagcggcaga acgtcactac gccaccatga gcatggagga   120
tttatgctca ctgcctatca gccgtttttgc tgcgccggac tgcgcgcttt ttctttggag   180
tacattccca cagcttccag aagcactgca tctgattcgt gcatggggat tcaccta   237

Seq ID 115
cgaaaacaat tgatcacagt cgtttacaag tcaaagattc aacgatttat gtcggcgatt   60
catggaaacc agaagagaac tttgtttcag caacagataa aacaggtcaa gacgttccct   120
tcgaaaaaat cactgtttca ggtcaagttg ataacaccaa agcaggcgtt tatccaatta   180
tttacagtga cgaaggtaaa gaagaaacag cctatg   216

Seq ID 116
caggaattac ggatagtggt tatacagata caaataaacc agaaacggat acgccagcaa   60
cagatgcagg cgaagaaatt gaaaaaatac ctaattctga tgttaaagtt ggcgataccg   120
tcaaag   126

Seq ID 117
ttatcatggc aaagcagtgc aagcagaaga tgttttggaa atcgtgacaa atcaataccg   60
agcagtcggt ggcggaaact atcaaatgtt ccaacc   96

Seq ID 118
agaaacacgc aagggaaatg ccgatttttcc ttgcgtgttt cgctgttttt cggtaaagta   60
tgtagaaaag tagggagttt tgctcatgaa ttatccattg atttttgcag aaaaccaacg   120
aatcgaaacc gaacgcttga ttttaaggcc cgtcacatta gccgatgcag aggatatgta   180
cgaatatgct tcaga   195

Seq ID 119
tttcccagcg gacttctttt cgccagagga aacgttggcg ctggttccgc aagcggtgat   60
tcaagcattc gcggaacagt tgaacacgcc acttttaacg gcagccgaac aagcggcggc   120
accgaatctg ctagaagcat t   141

Seq ID 120
aaccaacgga agaaattcaa ccaaaaccag aatcaatcac ccctcaggaa aagtcagaac   60
cggaacaacc tgcgccaacc cctccggaag atcaaaaaac tacggttcct tctgaagaaa   120
aaaagccttc acccaagaaa aaagacaaac ctgtggagaa acggcaatg tcaaaagaaa   180
aaaagaaaca aacaaaagaa acaacggctc atcccttaca ag   222

Seq ID 121
caacaagttt gtttggcgcc tttgtcgcac tccttgcgag catgggaatc atcaaacgca   60
aacgtaaaaa ctagctctct cctt   84

Seq ID 122
gctgttatgc acgaccaaaa gacaaaccga cagcccttga cccaatcaac gaagaagaag   60
ccaaagaaca agcaaaatat gcggtcaatg gtttccctca agatttaca gaatggtacg   120
agtgggaaat taacaatgga aaacttaaaa atttactata aaaaaagaaa ccttttgcta   180
acgacaaaag atttc   195

Seq ID 123
ccaaactctt gtgtatacca catgaatgac caacgaatta agttccaagc aacgaagaat   60
aagattgggc ctaagatgtt accacccatt gctaaggaag cacctaaagc gcctaacatt   120
ggacgaacag taaaccagaa aactgggtca ccgacaccgg ctaaaggacc catcatccca   180
actttaaccc ct   192

Seq ID 124
attaaatgtc cgaccagccg tcagctgaag catttctttt ttcatgcctt ccgccaaacg   60
tgcatcccca aagctttcaa tcatatcaat aacggcgatt acttctttat aatatctcat   120
tttgccagtg tacagataat acgtaaggtt ttt   153

Seq ID 125
cgccatatca gtaatggggg caccttcctt gcggttactg aaatggccat tgatgacttt   60
gatagccgaa ctttccaaat tttc   84

Seq ID 126
tccacacttt caataatcgc tttaatattt tctaaacatt gcgtcgtttg ttccttgatt   60
ccgccagcga ctaacgcatt tgttttaggg tccaatggta gc   102

Seq ID 127
aacaagtcca gtgctttttc accagctaag cgatgtcccc ctttaaccac gacatgcttg   60
gctcccattt gtacaattcg ttttgccgct tcttccatct ccgccacgga agaaatttcg   120
cctaaaccag ataagatgcc cgcttcaatt aaattaggcg tggccactaa tgctaatggc   180
agtaaatcgt tttttaggcc ttccacactt ttgggttgca gaatttgtgc cgttccctta   240
caagcaatga ctgggtcaat cacgactttt   270

Seq ID 128
ttgccgttgg ttgtttataa tgaatcaaat cgcaaaatgg atgcacttga aaactcgtgc   60
cgacaatcac gattaagtcg gcagaagcaa cggcctgaat ggcttttttca attgcctctt   120
cggatagtcc ttcttcatac aaag   144

Seq ID 129
accagtaatg acagcatcat gattgctaaa ctaagcaaca gttttttcaa aaagtttcgt   60
tctcgaacaa aacaattaat cgctgcatcc gattgtcgga ctaaaacgaa taacgtcgct   120
agtaacccaa gaatagtaga ccaactagca cctaatgttc cccaacgaat ggtgaaaaag   180
cccgttgtca gtaacttcac taataagccc acaccagctg ccactaaagg acctttaaaa   240
cgattgcgac tttgttcaat actttgg   267

Seq ID 130
tcagctttaa tgtctactac acggccttct ggtttatcta ataaggtaac aattttcaca   60
gatgcagctt tacggtatct aaacaaatct actaaatacg ctaatgtacg accactatcg   120
atgatatctt ctacgattaa aatatgtcta ccttcgacat ttgtgtccaa gtctttgaca   180
atttttactt cgcctgaaga aactgtcgcg ttgccg   216

Seq ID 131
cacgtgttac agaaataccc acgcccgcaa tacctacgct atccgcggaa aattgagcca   60
tttccccagt ttgtaatttg ctttctggaa tttcaaccaa tacaaaacct t   111

Seq ID 132
atgtcaataa tcacaccagc agataccttt ttgaatcgtt taatcaacgt agctgctgac   60
tctttcaagc catttctttc taacgttcct gttgctaaca taataataat aaagatagcc   120
atgcctcgat tattcacaaa ggttgttcct agtgtttcta gcattgatgt caaaccaagc   180
ccgccagtaa aagccgtcac tattaaagca accatcacaa tcaagatagc atcaaatttc   240
aatgcgaagc cgacaatgac gatcaatacg cctaataaag gtaataaatc catcaaatct   300
ctccttctta aaataaaaatt tcacttttcca ttatatgcca attttcgctc atttgtaaga   360
cttttt   366

Seq ID 133
cggctaaata gatcggttct ttcgttggtt cctctgcttg cagtattagg cgataaacaa   60
aaaaatcacc ttgcgttttt tgtaaaagct gctgctgtaa gtagcttaat agttcaccat   120

Seq ID 134
agtgtgtgta ttgctcgcca cctgccacat ttggatgaac tcggtcatca taaaaccaat   60
catcatgggc gttactataa t   81

Seq ID 135
gttgcaccaa acaaggtatc aggacgggtt gtaaacaccg tgaaactttc ttctgtgcca   60
gcgactttaa aggtcacatt ggctccttca gaacgtccaa tccaatttcg ttgcatatct   120
ttaatactct ctggccaatc aacaagctct aaatcttcta ataagcgatc tgca   174

Seq ID 136
atgattcctg catccataaa aatgacacgg tccgcgactt cttttgcgaa gcccatttcg   60
tgggtaacca ctaccattgt catcccttca atagctaaat ttttccccta ggatagcact   120
tcgcccacca tttcagggtc taaggcagaa gttggttcat cgaacagcat cacgtcaggg   180
cgcatggcta aagcacgagc aatcgctacc cgttgttgtt gcccaccaga aagt   234

Seq ID 137
cataatccca tgaaaattga agcaaaaaat gctggaatcg catcttctaa atccgtccaa   60
ttgatgtctt tgaaagaagc taacatcata attcccacta aaattaacgc tggagccgta   120
gcttgtgctg gcacaatagc aataagtggt gaaaaaagac tacttagggc aaatagaatt   180
gctaccacta ccgatgtcaa acctgttcga ccacccgcac cgattcctgc tgcactttct   240
aca   243

Seq ID 138
tattttttcg taatttctac gaagtattca tctaacaaca agtttttcaag tgctggattt   60
ttttcatatg catcggtaat ttttttgtaag aattgtgcac ggataatiaca acctgctcgc   120
cagattttttg caatttcacc aaatggtaaa tcccatccgt attcttcaga agctgcacgg   180
agttgtgcaa aaccttgcgc a   201

Seq ID 139
ccatcttcat caatcaatgt gacttttgtc gttgtggagc ctgcatcaat ccctaaaaag   60
gtcacgccag tatgt   75

Seq ID 140
gcaataataa acgatgtccg tcctttcaga agttgttgta gtccttcttg taacaattct   60
tctgtttttg tatcaatact ggaggttgct tcgtctaaaa ttaaaatttt gggatctgcc   120
agtaaagcac gagcaaatga aattaattgg cgttgtcctg ccgagagtgt actaccccgc   180
tcttccacaa ctgtttca   198

Seq ID 141
tcttcctctt ctggaaacac ttgtggccaa aaagattcag ctagccaaaa taattgctct   60
gttacagttc ctttaaactt aggaagttcg tcttcagcca agggcaaact cgcaacttca   120
gctgccgta   129

Seq ID 142
atcatccccg cttcaacagg tgaatgtaaa gcatcgatgc caaacatagt taaggccaat   60
gcaatcgcca acgtatacaa caaattaatg atgggactta ataataacga attcatacga   120
atcattgcta ctcgtgtttt taaa   144

Seq ID 143
actaattgca ttgctacgcc acaagcgatt cctaatttta aatctcctaa aataatgcca   60
caaactaatc cgccgactaa aggacgccca atcgta   96

Seq ID 144
tacccattcc aaatagacaa atcaacgcct tggtctcctt ttgcagcaaa tgtgctaatt   60
ggtgatacaa aaaataagac taccagtaac gttgccaata gtttcttttt c   111

Seq ID 145
ttttttaatg atttaagaat atccataatc actttttcac tagacatatc aatgccgacg   60

aaaggttcat ctaaaaaaat caccctcagca ccttgggcca aaacacgggc gataaacaca   120
cgttgtaatt ggccaccaga aagttcgcca atctggcgtt gcgca   165

Seq ID 146
tttttacaca agagtcattg tcctgtagaa ccgaaaaaaa cagttcctgt gccttcgtta   60
atttgt   66

Seq ID 147
gacacaaatt gctgtgcaaa cagtgttgcc gttacaccga caagcagtaa caggacaatg   60
atcgta   66

Seq ID 148
gggtcgccgt catcgtcata ccgggtttcc acgatcacct cacgggtgag ggtgtactgc   60
cgggaaaaaa tgcggtcaat ttcgggctgg acttctgcaa aggtaaaagc attgaacgcc   120
gtggagagg   129

Seq ID 149
ttgctgtgcc actattgttt cttcctgatt ccacagtgtg atagaaacaa cagggacagt   60
ttgtgccacg tccttggttc c   81

Seq ID 150
aagacaatcg caaaggttaa taatgcaact aaccaattcc gggctgaacc aaaatcagca   60
ctaccagctc ctcctgccat atacttaatc gcaacaggaa acaaggatag tccaatgctt   120
aaaataaccg ttcctgttac taaaggcgga aacaacacac ggatttttttt aatgaacaaa   180
ccaacaataa taacaagaat actgccaact aattgggaac caaaaattgc tgcaatccca   240
aaatcagctc cgatggccat caaaatcggt acatacgcaa aacttgcccc catcataact   300
ggtaaccgtg aacccacttt tccaaaaata gggaaaagt   339

Seq ID 151
ccaacactca tattaaaaat gaaaccattc gggtcaccta gctcgaactc tttactaagt   60
aacttcaaaa cgaaccatgc tttgacgtat tcatca   96

Seq ID 152
ggaacactat caatcacagt tgcaccagta tgcaccatga ctgctgctga aagtgaagat   60
gtgcccatat ctaaaatcgc tggtccaaaa gaacctgtcg ctaaaataac cgctgtggat   120
gttgaaccaa ctgctaagcc cattaaaata ccagaaaatg gtgccaagaa ggttccagaa   180
attcccataa cagaaattaa gtggacaatt tgtgttgata aatcagaaac agaaattaat   240
ccagcaattg cacctgcacc aattaaaatt aaaacggtat cggttactct cttcaagcca   300
gaagatgca   309

Seq ID 153
actaaatcac gtgctaacgt ggatacatta caagaaatat acaccatttt cttaggtggt 60
tgtttcaaaa gggccgttaa taatttacga tccaaaccag tccgtggggg atcaactaca 120
atggcatcag gtttaaagcc ttctttgaac catttaggta ataagtcttc agctgttcca 180
acagca 186

Seq ID 154
ctttcctgtg aaactaaaaa gtccacgggg ccatcactgt tgaccgtgtg catactaatc 60
gtatacacaa aaaccatcac gccgatccca ctcaaggctg tcgcaattaa taaggtcaaa 120
ccagtgagtg ttttgctact tttaaataag 150

Seq ID 155
tttctctttt cacttccggt aaggcgtgaa ccaccactaa aattccataa acaaagttcg 60
gtgctagccc cttttgttcc aataacaaca ataattcttc aaaatcaaca cctacagcaa 120
aagctagtcc taaaccggca aaggcatacg ttcggc 156

Seq ID 156
aataaaccag caaaaaacgc gccaaaaacg gcacctgtta acgctaatgg cggaatacct 60
aagaatgtca tacggattaa gccgcagaca agtgccatga gcgttccg 108

Seq ID 157
ataaattccc aaatagcgcc aatgcccata aattgcctca acctgtgcgc cgatataata 60
aagagtaacc atattcaaaa taatatgcat aaatccaatg tgaagaaaca ttggcgtgat 120
aaaccgccag tattcatgat tttgagctac cagtggacga accattgctc cccaattaac 180

Seq ID 158
ttttcgtttt cgctttacga attcccgcat aaaaagcggc aaataatcca acaggaccgc 60
cgccga 66

Seq ID 159
gcaaagaatg cattaattcc taaagctggt gctgtcgcaa tcggatattt cgccaacaca 60
cccattaaaa tacaacctaa ggcactcgct aaagcggtgg cagtaaaaac agcgccttcg 120
tccatgccag aagcaccaag tacggttgga ttaacgaaca aaatatacgc cattgagata 180
aacgtggtga aaccagctaa catttctcgt ttcatattcg tattcaattg ttcaagttca 240
aaatatgaac ttattttttc tttcatgttc cgtcctccta aaatc 285

Seq ID 160
ctttttttcag atacattctc tatttccctc accgcattcg tttctttaaa ttcaatcatt 60
ggaaatagcc gacgcaattc cgattgcaaa atcaacgaag aactgatacc actgggacaa 120
acaattagcg cttttagctt ttcttgccgc tgttcatttt tttgcccgaa aatttctcca 180
ccaaataata tggtaaaaaa accgatttcc tctgacggga ttgaccgttt cgttaatttt 240
tctaaaggag ctaacgcttg ttcacaaaga 270

Seq ID 161
ataaattccc atagcataca acgtgttatt cgggtcccat gtccaaaaag ctcgtctggt 60
tctaaaaaaa gttccagcgt cagt 84

Seq ID 162
gtaaatcaat gtctgaaacg acaaatttat cgccatattt ttttgataca tcactcatct 60
ggatcatgtg cgttgcttcc tttctgcaat aatttttccat ataaaataaa cgcccccaac 120
aaattgaatc accacactga gcggggtatt caaatgaaag acctgttcaa ccatcaattg 180
accaaacacc ag 192

Seq ID 163
tcggccatct caatttttc tgttccagag ctgatgtata aagtaacaga tcggcctttt 60
ttcacagatg attttgcttc ggggtctgtt ttgaccacct tgccttcttc aatcttatcg 120
tcgggaattt ttttggtttc actatcgact ttcagcccgg cacttttag cgcttgtgaa 180
gcgtccgctt tcgtttcatt tgtaacatca ggaacttcaa cgtctttacc gccgacatt 240
gcaaaggcta agccacctat tgctaaagaa attaatgccg caattgctaa gccaatccac 300
aatttttac ggtttttctt cggcggttgt tcagctaaaa tgtcatctgc gatatcttct 360
gggacttcga caggcgttgt ctcttccggt tcagcgatgt cttcgggaat cggagttaat 420
actttcgttt ctcctaataa agccgttggt tcccacgcag gttcgtttaa acgactggct 480
gacaaggacg tgtataagtc ttctgccatc tcattcgctg ttttg 525

Seq ID 164
ggatcaccgc cacctcctgt ccctaaaaaa gctgctcctg tggctatatt ctcaattgct 60
tctgcatcta aa 72

Seq ID 165
tctccagcaa ttactgcgcc taatgcaaaa gaaccgacag tcattgaaag acctgacata 60
aagcttcctg ctgtt 75

Seq ID 166
```
attcgttccc gtaatgatcg cttagctagc gggacactga tttccttctc tgtaaaatgc    60
aacactagtg caaacaaat  ggaaaagcca atgcctaata acatacagcc actcatgaac   120
ggaatttttg ctagcggaat gctcgtactg gcacccatcg ctgtcaccat cacaccaaag   180
aaggcccctg gttttgaaat atcg                                          204
```

Seq ID 167
```
atgtttgtaa acattggcat gaacccaatt ttcattgaca gaaaggcaat caagaccgat    60
acaatcggcg cgaccgcaat cggaatcaat aataaaggg  taaaaataac tgggagtcca   120
aaaatcaagg gttcattaat gttgaaaaat gctggaacaa acgaaatttt cgacatggct   180
ttcattctag atgaatgtcc gacgatggcc attgcaatag ccaagccaat cgtgcaacca   240
gcgcctccaa ttgtagcaaa t                                             261
```

Seq ID 168
```
caagccaact cggttcgtcc catgaattcc catgcgcacc tacaggcgga ttcaaataat    60
cccccttctgc acgtttttgga t                                            81
```

Seq ID 169
```
gtaaaattac cagtaccatc attcgaaata tcttgtgctt tttgttcaag acttcgcat     60
cctgataata gcaaagtggc tgctgttagc agtagccact tttttccaat aaatagtttc   120
```

Seq ID 170
```
gtgatccaac aagaaagtcg tttaagagtc gctgacaatt ctggcgcacg tgaaatttta    60
acgatcaaag tcctaggcgg atctggtcgt aaaactgcta acatcggtga cgttattgtt   120
gctactgtca aacaagcaac gccaggtggg gttgttaaaa aaggtgaagt agttaaagcc   180
gttatcgttc gtacaaaatc tggcgctcgt cgtgctgacg gttcttacat caaatttgat   240
gaaaatgctg cggtaattat ccgtgacgat aaaagcccac gcggaacacg tatcttcgga   300
ccagttgcac gtgaattacg tgaaaacaac ttcatgaaga tcgtttctct agcaccagaa   360
gtatta                                                              366
```

Seq ID 171
```
MVGIIILASHGEFAEGILQSGAMIFGEQENVKAVTLMPSEGPDDVKAKMQEAIASFDNQDEVLFLVDLWGGTPFNQANSLLE
DHKDKWAIVAGMNLPMVIEAYASRFSMESAQEIATHILETAKDGVKVKPEELQPAEAPKAAATEDAQPKGSLPPGTVVGDG
KIKFVLARIDSRLLHGQVATAWTKATQPNRIIVVSDAVAKDDLRKKLIEQAAPPGVKANVIPISKMIEVAKDPRFGNTKAL
LLFENPEDVLKVVEGGVEIPEVNVGSMAHSVGKVVVSKVLSMGQEDVDTFDELKAKGIKFDVRKVPNDSKANMDEILKKAK
NELANA
```

Seq ID 172
```
MKNKSRQFLKNNWPYMLASFFIPFLIMAIIYLSIGIYPGSSRSVLASDAFSQFSNFHASFNNVLHGKQSLFYTWNASLGLN
YLSLISYYLGGLFTPLVFPFNNQNMPDALYFLTLLKIGSAGLSFWFLAKQTFKIPKWSHVTLSVSYALMSFIVAHSELIMW
LDAFIYLPLVILGIHRLMDQRKPTLLFVSYFLLFITNYYFGFMIGLFSFLYYFARTFTDWQRYKSRIVAYFTTSLLAGGAS
MIMVLPAVLDLRTNGETLSEITTFKTEATAFLDIIMKNMIGVYDTTKYGSIPFIYIGLLPLIFCLFYFVTKEVPLKNKLLF
GSLFVLLIASFYITPLNLFWHGMHAPNMFLFRYSFLFSFLVILLAGYGWEKFEKDDLGVLSGLILILLAIFALAMGTKGAT
SYTYVTLTSFVLTATFLLLYFFGIAFYQLKKAPMQYLVILLLLLVSGEAIINTRAMVTGILDDWNYASRSLYSEPYPDLKN
LVDQTKKENDTFYRLENLNGVSANDGINYGYSGISMFSSVRNRHSSTYLNALGFRSRGTNLNIRYQNNTLLMDALMGIKYN
IAENNPMKFGPERQAAAGKYQLYRNENALPLGFLADKEIYNVRQPLNDNLGSQTNLLNALANTNERYFTFYQPTMTLQNNV
TITQNTAGVTFTEKQHNVAKEISYTVNVPANTQAYLSLFPTDFAQLESSTATVTVNGSSQQSQIGITGQYYNLGYYPKDTT
VNFKVSFYGTKAVSFVQPQVVGLNTNAFEKAISAVQEKGVDLTTGKRSASGTFTADKDQVLVTTIPYDKGWRVKIDGKKVT
PKAFKDAFLSVPVSAGTHTIQFSYLPEGLIPGIVLFVLCTGGFVAYVTLIPARRNRKKEDK
```

Seq ID 173
```
VKKLSFKKVKWGMHFLMAVALIAPSVTSTAYAVETTSQQSSEAVTSTTDSSRKQEPVITQETTDIKQEAPNQATSDSVKQS
QETTAPTETTNLETSIAEKEETSTPQKITILGTSDVHGQLWNWSYEDDKELPVGLSQVSTVVNQVRAQNPAGTVLIIDNGDN
IQGTILTDDLYNKAPLVNBKTHPMITAMNVMKYDAMVLGNHEPNFGLPLIKKIQQEATFPILSANTYNKEDGLRFVEGTTT
KELDFNQDGQPDLKVGIIGLTIPHIPLWDGPRVTSLNFLPLKEEAEKAVTELKANDQADIIVASIHAGQQNSDPAASADQV
IENVAGIDAYILGHDHLSFTKQGAAPNGKTVPVGGPKDTGTEVVKIDLSVAKNADKWEVQEGTATIVPTTNVPADEAVKAA
TKEYHEKTRAFIQEEIGTATADFLPKQEIKGIPEAQLQPTAMISLINNVQKEVTGAQLSAAALFKYDSKLPAGKISYATIF
DIYKYPNTLVSVPINGENLLKYLEKQGAYYNQTQPDDLTISFNPNIRVYNYDMISGVDYKIDISKPVGERIVDAKIDGQPL
DPAKEYTIAMNNYRYGGLASQGIQVGEPIKNSDPETLRGMIVDYIKKKGTLDPEQEIERNWSIIGTNFDKWRAKAIELVN
DGTLQIPTSPDGRTPNAAAITKQDVRNAGFDLDNAYTIMHTNDVHGRLEAGKGELGMARLKTFKDQENPTLMVDAGDVFQG
LPISNFSKGADMAKAMNEVGYDAMAVGNHEFDFGLEIALGYKDQLNFPILSSNTYYKDGSGRVFDPYTIVEKSGKKFAIVG
VTTPETATKTHPKNVEKVTFKDPIPEVEAVIKEIKEKYADIQAFVVTGHLGVDETTPHIWRGDTLABTLSQTYPELDITVI
DGHSHTAVESGKRYGKVIYAQTGNYLNNVGIVTAPESEPTKKTTKLISAAELLELPENPAVKAIVDEARTNFNAENEKVIV
DYIPFTLDGQRENVRTRETNLGNLIGDAIMSYGQDAFSQPADFAVTNGGGIRADIKQGPIKVGDVIAVLPFGNSIAQIQVT
GAQVKEMFEMSVRSIPQKDENGTILLDDAGQPKLGANGGFLHVSSSIRIHYDSTKPGTRLASDEGNETGQTIVGSRVLGIE
IKNRQTQKPEPLDEKKQYRMATNDFLAAGGDGYDMLGGEREEGISLDSVLIEYLKSATSLRLYRAATTIDLAQYKEPFPGE
RIVSISEEAYKELIGGGETPKPDPKPDPKPTPETPVATNKQNQAGARQSNPSVTEKKKYGGFLPKTGTETETLALYGLLFV
GLSSSGWYIYKRRNKAS
```

Seq ID 174
MNQQTEVKKRFKMYKAKKHWIVAPILFLGVLGAVGLATDNVQAAELDTQPGTTTVQPDNPNPQSRNETLKTAVSEEAALQK
DTTSQPTTAEEVVPKGIAAEQSSATSNDTTNVQQPTAEAEKSAQEQPVVSPETTIEPLGQPTEVAPAENDANKSTSIPKEF
ETPDVDKAVDEAKKDPNITVVEKPTEDLGNVSSKDLAAKEKEVDQLQKEQAQKIAQQAAELKAKNEKIAKENAEIAAKNKA
EKERYEKEVAEYNKHKNDKGYVNEAISKDLVFDSSIVTKDTKIDKITGGKFIKASDFNKVNQGQSKDIFTKLSKDMNGKAT
GNFQSSKVAAVEFGPKGGYAVLLEKNKPVNVTYTGLNASYLDRKITKAEFIYELQSAPSQSGTLNAVFSNDPIITAFVGTK
NANGKDVKVRLTIKLYDANGKEVLPDKDHAFAYALSSLNSSLGTNYSVEHAEFVSDFGSKNEFKYINGSYVKKQADGKFYS
TEDLDYGTGPSGLKNSDWDAVGHKNAYYGSGVGLAREGGRISFSFGMTTKGKVNLSGAQWFAFSTNLNAKSIRPYQKKGNP
KEPEKATIEFNRYKANVVPVLVPNKEVTDGQKNINDLNVKRGDSLQYIVTGDTTELAKVDPKTVTKQGIRDTFDAEKVTID
LSKVKVYQADASLNEKDLKAVAAAINSGKAIDVTASYVLNLDQNTVTAMMKTNADGSVVLAMGYKYLLVLPFVVKNVEGDF
ENTAVQLTNDGETVTNTVINHVPVSNPSKDVKADKNGTVGSVSLHDKDIPLQTKIYYEVKSSERPANYGGITEEWGMNDVL
DTTHDRFTGKWHAITNYDLKVGDKTLKAGTDISAYILLENKDNKDLTFTMNQALLAALNEGSNKVGKQAWSVYLEVERIKT
GDVENTQTENYNKELVRSNTVVTHTPDDPKPTKAVHNKKGEDINHGKVARGDVLSYEMTWDLKGYDKDFAFDTVDLATGVS
PFDDYDETKVTPIKDLLRVKDSKGADITNQFTISWDDAKGTVTISAKDPQAFILAYGGQELRVTLPTKVKANVSGDVYNSA
BQNTFGQRIKTNTVVNHIPKVNPKKDVVIKVGDKQSQNGATIKLGEKFFYEFTSSDIPAEYAGVVEEWSISDKLDVKHDKF
SGQWSVFANSNFVLADGTKVNKGDDISKLFTMTFEQGVVKITASQAFLDAMNLKENKHVAHSWKAFIGVERIAAGDVYNTI
RESFNNEKIKTNTVVTHTPEKPQTPPPEKTVIVPPTPKTPQAPVEPLVVEKASVAPELPHTGEKENTLLSVLGAGMLVGLAW
FGLKKREVK

Seq ID 175
MKRIGYARTTIIEDDLKNQLTTLQSFGCDDIFQETFDPQAEISVLDEVEKLLSAGDTLIVCKLHHLGKTTRQLTDFMKMLK
EKQVDFVSISEGIDTHLPTGEAYFQLMESLSAMECALIKERTLVGLHKARENGKVGGRPKIDGRTVRKIRALYYENKETIQ
FISNKCGVSVGTCYKYINLPETDVERLYS

Seq ID 176
MSKKEINQVVASSYQLYINGEWTTGSGNKMIASYNPSNGEKLAEFVDATNADVDRAVEAAQEAFQTWKDVDVVTRSNLLLK
IADLIERNQEHLAMVETLDNGKPLRETQSIDVPASADHFRYFASVIRGEEGSVKEFDKDTLSIVVKEPIGVVGQIIPWNFP
LLMGAWKLAPALAAGNTVVIHPSSSTSLSLLELFKIFDQVLPKGVVNLITGRGSDSGNYMLAHPGFDKLAFTGSTEVGYTV
AKAAADRLIPATLELGGKSANIIFEDANWERALEGVQLGILFNQGQVCCAGSRVFVQSGIYDQFVEALKEKFEQVNVGFPW
EKDVEMGAQINEHQLEEILKYVEIGVKEGATLITGGQRLTENGLDKGAFLAPTLLANGTNAMCVAQEEIFPGPVATVIKFET
EEEVIRLANDSEYGLGGAVFSQDINVALRVARGVRTGRMWVNTYNQLPAGAPFGGYKKSGIGRETHKSMLDAYTQMKNIYI
VTKEEADGLY

Seq ID 177
MGKYHELAEKIVKNVGGQENINSLTHCITRLRFKLKDESQANDDVLKNMDGVVTVMKSGGQYQVVIGNHVPAVYEEVVSIA
GLSGERREEASSGNLFDRLIDILSGCFQPFLGALAAAGMVKGLNALLVFLKLYTATSGTYTMLNGIGDAIFYFMPVILGYT
AAKKFRLHPMVGIVIGAALCYPTIQGSALQTAFETTAGAGAAAPYNLFGLPAYNTFMGIPWVGANYTSSVVPIIFIIAFAA
QVQKVFKRIIPEVVQTFLVPFFVLLIALPIGFLVIGPIVSMLTDLLSAGFTALMSFSPALYGLILGFFWQVLVIFGLHWSV
VPLAIMQVTQEGSSQVLTGSFAASFAQTAVVLAMFFPKLKDKKLKALCPPAIISGIFGVTEPAIYGITLPKKWPFIYSMIGG
AVGGLYLMINNVTAYTMGGLGIFGVLNFINGDDASGMIQSFIAIAIAAVVGFGLTPFFWKDNTVEEEEVIIDKTTIKKENI
TSPVKGRVLSLKNAEDPAFANGALGNGVVIEPTEGKVVAPFDGTIVTLFPTKHALGLISDNGTELLIHIGIDTVQLEGEGF
EAFVKQGDRVKKGQTLVTFDLEGIKKAGFSTQIPIVVTNTADYLDILEVGSNEVSTSDDLLTALI

Seq ID 178
MESKTFDIEGMSCASCAQTIEKATAKLPGMAKASVNLATEKLSVTYDQTEVTEEEIKEAVSDAGYKAISPAQQRTFAIEGM
SCASCAQTIEKAVNQLSGVQQAIVNLATEKLVVSYDDHQVTSAEIIKAVTDAGYQATEEVAAGATADQDREKKQKHIAEMW
QRFWISAVFTVPLLYIAMGHMVGLFLPDFLNPMTHAMTPAMVQLILTLPVLYVGREFFTVGFKALFKGHPNMFSLVALGTS
AAFVYSLYGTVMIFLGDTSFTMALYYESAGVILTLITLGKYFEAVSKGKTSDAIKKLMGLAPKTAHILRDGVEIEVPVDAV
QLDDIVIVRPGDKIPVDGVIVSGSSSVDEAMLTGESLPVEKKVGDAVIGASINKNGSFQFKATKVGKETALAQIIQLVEDA
QGSKAPIAQLADKISGVFVPIVIGLAVLSGLAWFFLGQESWIFALTITISVLVIACPCALGLATPTAIMVGTGKGAENGVL
IKSGDALETTHKIQTIVFDKTGTITEGKPVVTDILVADSALSEAELLTLAASAEQGSEHPLGEAIVGAAKERQLPLAEGSD
FSAIPGHGIRVTVNERVLLLGNIKLMKEEAIELSTFVQQADRLAEEGKTPMFVAKDGSFAGIIAVADTVKDSSQTAIARLH
KMGIEAVMITGDNKRTAEAIAKQVGIDRVLSEVLPEDKALEVKKLQAEGKKVAMVGDGINDAPALAQADVGIAIGSGTDVA
MESADIVLMRSDLMDVPTAVELSKATIKNIKENLFWAFAYNTLGIPVAMGVLHLFGGPLLSPMIAAAAMSFSSVSVLLNAL
RLKGFKPSTVKRTSGSQK

Seq ID 179
MKKKILAGALVALFFMPTAMFAAKGDQGVDWAIYQGEQGRFGYAHDKFAIAQIGGYNASGIYEQYTYKTQVASAIAQGKRA
HTYIWYDTWGNMDIAKTTMDYFLPRIQTPKNSIVALDFEHGALASVPDGYGGYVSSDAEKAANTETILYGMRRIKQAGYTP
MYYSYKPFTLNHVNYQQIIKEFPNSLWIAAYPIDGVSPYPLYAYFPSMDGIGIWQFTSAYIAGGLDGNVDLTGITDSGYTD
TNKPETDTPATDAGEEIEKIPNSDVKVGDTVKVKFNVDAWATGEAIPQWVKGNSYKVQEVTGSRVLLEGILSWISKGDIEL
LPDATVVPDKQPEATHVVQYGETLSSIAYQYGTDYQTLAALNGLANPNLIYPGQVLKVNGSATSNVYTVKYGDNLSSIAAK
LGTTYQALAALNGLANPNLIYPGQTLNY

Seq ID 180
LPMIIKQEQFIPKDTVETTIDLLIRNLTTIKDNTGEFLLDFDGLKVDDKSWTIWNWPQGVGLYGIYKNYRNTKSEKALQVV
NDWFEGRMQEGAPPKNVNTMAPLLTMAYLYEDTKDSKYIPYLEQWAEWVMEEMPRTNEGGLQHATYGPENKNQLWDDTLMM
TVLPLAKIGKLLNRLDYLEEAKHQFLIHIKYLQDKKSGLWYHGWTFEGNHNYAEALWARGNCWITIAIPEIIEILELPKGD
SLREFLLSTLNAQVAALAKYQDESGLWHTLINDSNSYLESSATAGFAYGILKAVHKKYISSEYEEVANKAIAGLLNEIDET
GEVQHVSVGTGMGDNLDFYRTIGMTAMPYGQSLTILCLTELLVSYC

Seq ID 181
MKAKKQYKTYKAKNHWVTVPILFLSVLGAVGLATDNVQAAELDTQPETTTVQPNNPDLQSEKETPKTAVSEEATVQKDTTS
QPTKVEEVAPENKGTEQSSATPNDTTNAQQPTVGAEKSAQEQPVVSPETTNEPLGQPTEVAPAENEVNKSTSIPKEFETPD
VDKAVDEVKKDPNITVVEKPAEDLGNVSSKDLAAKEKEVDQLQKEQAKKIAQQAAELKAKNEKIAKENAEIAAKNKABKER
YEKEVAEYNKHRNENSYVNEAISKNLVFDQSVVTKDTKISSIKGGKFIKATDFNKVNAGDSKDIFTKLRKDMGGKVTGNFQ
NSFVKEANLGSNGGYAVLLEKNKPVTVTYTGLNASYLGRKITKAEFVYELQSSPSQSGTLNAVFSNDPIITAFIGTNRVNG
KDVKTRLTIKFFDASGKEVLPDKDSPFAYALSSLNSSLTNKGGHAEFVSDFGANNAFKYINGSYVKKQADGKFYSPEDIDY
GTGPSGLKNSDWDAVGHKNAYFGSGVGLANGRISFSFGMTTKGKSNVPVSSAQWFAFSTNLNAQSVKPIFNYGNPKEPEKA
TIEFNRYKANVVPVLVPNKEVTDGGQKNVNDLNVKRGDSLQYIVTGDTTELAKVDPKTVTKQGIRDTFDAEKVTIDLSKVKV
YQADASLNEKDSKAVAAAINSGRAKDVTASYDLHLDQNTVTAMMKTNADGSIVLAMGYKYLLVLPFVVKNVEGDFENTAVQ
LTNDGETVTNTVINHVPGSNPSKDVKADKNGTVGSVSLHDKDIPLQTKIYYEVKSSERPANYGGITKEWGMNDVLDTTHDR
FTGKWHAITNYDLKVGDKTLKAGTDISAYILLENKDNKDLTFTMNQALLAALNEGSNKVGKQAWSVYLEVERIKTGDVENT
QTENYNKELMRSNTVVTHTPDDPKPTKAVHNKKGEDINHGKVARGDVLSYEMTWDLKGYDKDFAFDTVDLATGVSFFDDYD
ETKVTPIKDLLRVKDSKGVDITNQFTISWDDAKGTVTISAKDPQAFILAHGGQELRVTLPTKVKADVSGDIYNSAEQNTFG
QRIKTNTVVNHIPKVNPKKDVVIKVGDKQSQNGATIKLGEKFFYEFTSSDIPAEYAGVVEEWSISDKLDVKHDKFSGQWSV
FANSNFVLADGTKVNKGDDISKLFTMTFEQGVVKITASQAFLDAMNLKENKNVAHSWKAFIGVERIAAGDVYNTIEESFNN
ETIKTNTVVTHTPEKPQTPPPEKTVIVPPTPKTPQAPVEPLVVEKASVEPELPQTGEKQNVLLTVAGSLAAMLGLAGLGFKR
RKETK

Seq ID 182
MKKILFASLFSATLLFGGSEISAFAQEIIPDDTTTPPIEVPTEPSTPEKPTDPTPPIEPPVDPVEPPITPTEPTEPTT
PTEPTTPTEPSEPEQPTEPSKPVEPEKPVTPSKPAEPEKTVTPTKPTESEKPVQPAEPSKPIDVVVTPTGELNHAGNGTQQ
PTVPIETSNLAEITHVPSVTTPITTTDGENIVAVEKGVPLTQTAEGLKPIQSSYKVLPSGNVEVKGKDGKMKVLPYTGEEM
NIFLSAVGGILSVVSGFVIFKKRKAKV

Seq ID 183
MGSRFGLMLKQGGNDVLLIDGWQEHINAIKENGLKANYNGEEITVKVSIVNQNEVPTGEQFDLIILFTKAMQLEKMLQDVK
PLIADHTEVLCLLNGIGHEDVIEKFVPMEKIFIGNTMWTAGLEGPGKAKLFGSGSVELQNLGIGQEESAKKLAETLSASGL
NAKYSDNIHYSIYRKACVNGTMNGLCTILDVNMAGLGATKPAHDMVVTIVNEFAAVAAKENVNLDIPEVIEHVETCFDPTT
IGMHFPSMHQDLIKNNRLTEIDYINGAISRKGKKYGVVTPYCDFLTQLVHSKEEILGAK

Seq ID 184
MQTTEEMQSGDGRRGRLRIFFGFAAGVGKTYGMLTEAHELLMMGKHVVVGYIEPHDRPDTNRLLEGLPQIPPKNIFYKQMV
LTEPDIDQIIQQKPEIVLIDELAHSNAEGSRNRKRYQDIDELLNAGIDVFTTVNVQHIESLNDIVEEVTGIEVKETVPDTF
LRQATIRVIDVEPDELIERLEQGKIYANENAKRALKNFFIPQKLDQLRGLAIQRASDHINRISGKTIGIQSKLLTVVNDAF
PKMTEKCIRWTARLAQGLVVDWTVIQVRTQENTPTNIPLADKLGAEVISIEEDDSFETIVEFAKMTGVTDIIMGKNLRQPW
YEKIFIEAFDDRLLKRLKDTELHLIPFNEEKRSLFFKTRKVIEGGGKDLVIAIGGVFLATVVTELMQYIHVGDQNLMLIYI
SFVLLVARTTSGYFWSSLSSILSVLSFNWFFVEPLYSLTVYKQGYPFTLLLMLVVALMSSNLMIRLKKQADTSMKKEHQME
ILYELNKRYVLVESRKQILDISATYLSRLLEREVIIFDRQVKTESVHCINEKKSILNNEDEAAVAFWAAKNQKEAGNGTDT
LNGAKGFYLPIAAGRKTLAVLGIERNADLDLENDQLNYLKLVLTQIAVILEQTELKDEKEQVELENEREKVRSNLLRAVSH
DLRTPLTVISGIAETLGIGNDLKEETQRKLLKDIQEESQWLIRMVENLLSITRINMDTMKVNKTAEPVEEVIEAVYKHLRK
VYPDGQVDIHLPEEVIFIQADPILIEQALFNLIENAFRHGENDLPVKLNVYQEKEQTVFEIENHGEIPLKQFQKIQSNLSG
TNEVPVDSKNGLGIGLSIVKTIVHAHNGKMEMTIGRGKTLVRIYLK

Seq ID 185
MKNFYKKKFALTDQGAEALTKASISSFPVYCINMVPAFIIMMLIDELVLENAKPRWLYFAVSFVTLLFMYWLLDREYENLY
NSTYKESAHLRVQIADDLSNLPLSYFSKHNLSDLSQTIMSDVEGIEHAMSHAIPKSGGMALFPPFISVMLLVGNVKMGLAV
ILPTLFSFVLILLSKKSQTKANTKYYDTLRENSEBFQETIELQQEINSFNLSKKVQDRLFKKMEESERIHLKVELSTFSVM
ALSSIFSYVSLAVVILVGVHLLLTGEVTILYVVGYLLAAIKIKDSFDSMKEAVLEIFYLAPKIQRIRAMKETSIQEGSDSP
LKSFDVELRDVSFSYDNNTPILDHISFTAKQGEVTALVGASGSGKTSILKLVSRLYDYDEGCILIDGYDIKRVSPASLFSK
IAIVFQEVTLFNTSILENIRIGNSQASDEEVKKAARLANCEDFIEKLPDGYHTLVGENGSSLSGGERQRLSIARAFLKNAP
ILILDEITASLDAENEKRIQESLNRLIQDKTVLIISHRLKSIEKVNKIVVMDQGKVVDQGTHSELYRRSEIYKNLIKKTKL
SEKFVYEKEAQSR

Seq ID 186
LTLSDTLGQREVTALSYDSRDVTAETLFFCKGLNPKEIYLENAVKDGLEIYVSEVPYEVPAQLGIIVTDIKKAMAVLSMAF
YDYPQNKLKLIGFTGTKGKTTAAYFTKYILDVATQQKTALLSTMNSTLDGKTFFKSALTTPESLDLYRMMATAVANGMTHF
IMEVSSQAYKTNRVYKLFFDVGIFLNITPDHISPIEHPTFDDYFYCKRQLITHSKVIVLNHEADYFPLLKETAQQQKVPAI
VYGSQPAPEVDYSFAVSSEDSLRFIVESPADALGLAGSYHLRLGGDFNKGNALSAAIASVLVGASKEECQQGIAATTVPGR
MESLTNTNGATVYVDYAHNYDSLKNLLTFVREEHPDGRLIVLVGSTGDKAISRRKDFGRVLSELADVAVLTTDDPASEDPA
KICQEIQAHITKEMPVYTVLDRGEAIAHALSLSTTADDAIVLAGKGADLYQKVNGVDEPYAGDFALAEAFINKKN

Seq ID 187
MFLGFNEMKYSKGRYVLVVLVMVLIAWLIFILSGLANGLAQGNRLAVDQWQANQVVLSKEANSNLNVSVLDENVKETISGG
KIAPIGQQSLAIRPADDKKAELTNVSLFGIEKESFLMPKVIEGNAFTDKNQVIASETLKNQGFKIGDKLTAGKYDEQLEIV
GFISKSSYNIVPVIYTSLDTWRSIKYGDNPAMAKMVNGFIVRSKDNTEVKTTNKDSQVLSISDFIEKLPGYSAQNLTLDGM
IYFLIVIAAFIIGIPIFVMTLQKTAMFGVLKVQGVPTSFLAKAVMLQTALLAVLGVAIGLALTGITVLFLPEAMPYATNGP
RMILFSVLLILSALIGGAFSIRTIAKIDPLIAIGG

Seq ID 188
MQIFPYSDDPNKRYHTWNYALRQQFGEKIFKVPIDGGFDCPNRDGTVAKGGCTFCSVSGSGDMIVAPSDPLPLQFQKEIQL
MHQKWPTVDQYIVYFQNFTNTHAPVDVIRHRFEQVVNEKGVVGLSIGTRPDCLPDEVVNYLAELNERFYLWVELGLQTTFE
ETSAAINRAHDYQTYLDGVAKLRKHGIRVCTHLINGLPGETPAMMRENVRRTIQDSDIQGIKLHLLHLMTNTKMMRDYNEG
RLQLMSKEAYVSVICDQLEMIPPEIVIHRLTGDAPFETIIGPMWSLKKWEVLNAIDAEMKRRNSYQGKYTVISGKEVFN

Seq ID 189
MKKESMSRIERRKAQQRKKTPVQWKKSTTLFSSALIVSSVGTPVALLPVTAEATEEQPTNAEVAQAPTTETGLVETPTTET
TPGTTEQPTTDSSTTTESTTESSKETPTTPSTEQPTADSTTPVESGTTDSSVAEITPVAPSATESEAAPAVTPDDEVKVPE
ARVASAQTFSALSPTQSPSEFIAELARCAQPIAQANDLYASVMMAQAIVESGWGASTLSKAPNYNLFGIKGSYNGQSVYMD
TWEYLNGKWLVKKEPFRKYPSYMESFQDNAHVLKTTSFQAGVYYYAGAWKSNTSSYRDATAWLTGRYATDPSYNAKLNNVI
TAYNLTQYDTPSSGGNTGGGTVNPGTGGSNNQSGTNTYYTVKSGDTLNKIAAQYGVSVANLRSWNGISGDLIFVGQKLIVK
KGASGNTGGSGSGGSNNNQSGTNTYYTVKSGDTLNKIAAQYGVSVANLRSWNGISGDLIFVGQKLIVKKGASGNTGGSNNG
GSNNNQSGTNTYYTIKSGDTLNKIAAQYGVSVANLRSWNGISGDLIFAGQKIIVKKGTSGNTGGSSNGGSNNNQSGTNTYY
TIKSGDTLNKISAQFGVSVANLQAWNNISGSLIFAGQKIIVEKGANSGSTNTNKPTNNGGGATTSYTIKSGDTLNKISAQF
GVSVANLRSWNGIKGDLIFAGQTIIVKKGASAGGNASSTNSASGKRHTVKSGDSLWGLSMQYGISIQKIKQLNGLSGDTIY
IGQTLKVG

Seq ID 190
MMIVVLLMSSCLFLFVSYQLYLRLLVEKLDSEDRTLKRKINFYKYQKNSRLMIYLLFSVCLLSVLLLGIVYTYYQLNQRNI
RMEQRIERLAQGQATQGDKIKKQAIKKTALNQFPWKQAVSAESAAVLTNYELQLAREWRPYLGETSITMIRSEKTQTLTLS
VFSVGLSYHEFQTGQDNIVALIAALNSVKEITMIDFNFTYRDQEQTLVKSIDTYARETLETNLEPVVIS

Seq ID 191
MSNYTIIDLNQQPHLIPSAAQWFSQKWQIPAEAYQASMQEMLTSENHVPHWWVCLNKQQDIVAGVGVIDNDFHNRADLTPN
LCALYVEPTFRHQGLAGILLTTVGDFLANSGFKKLYLLTDHTTFYERYDWEFLTMVTTEEQSCARIYQKQLRQ

Seq ID 192
MKNSENDYIQSLFQILPGLLTAFLVACLSKFLAIWLPSLGAATIAILLGIFLGNTFVRGANLNRGTKVAESKLLEFSVVLL
GTTVTFQTIAQIGLQGVAFILIQMSLTIIFAYLIGKKLAFSDNMSLLMAGGNAVCGSSAIASIAPAIQADEEEKGQIITLV
NLLGTVLMLTLPILSGILYGTNLLARSALIGGTLQSVGQVVASANMVNENAVQLAMLFKIMRIVLLVAVVYLFGRFKQSKT
AESEAELVEVTKKSSALPWYVVGFFIACVFNSLIHFPVVISETAHFFSSWFEITALAAIGLRLDFKKFFQEGKRFLIYGLS
VGTVQVVLAILLLALLQF

Seq ID 193
MAKTGMYVGLDIGTTSVKVVVAEYIEGQMNIIGVGNAKSDGLNRGIVVDIDQTVQAIQRAVRQAEEKAGIQIKSVNVGLPA
NLLEVESCQGMIAVSSESKEITDEDVRNVASAALVRSTPPERQIVAILPQDFTVDGFEGIKDPRGMLGVRMEMFGVVYTGP
KTIIHNIRKCVEKAGLGINELVITPLALTETILTDGEKDFGTIVIDMGGGQTFTSVIHDKQLKFTHVNQEGGEFITKDISI
VLNTSFNNAEALKINYGDAYPERTSANEEFPVDVIGKSEPVRVDERYLSEIIEARVEQILRKSKEVLDEIDAFELPGGVVL
TGGAASMPGIVDLAQEIFEANVKLYVPNHMGLRNPVFANVISIVEYSAQLNDIYHIAKYAIPGEKSKPAQSVAVQQ
EVRYDTYAEQPQEEYEEFNERESGEKVTGKIKDFFSNIFD

Seq ID 194
MAKKTIMLVCSAGMSTSLLVTKMQKAAEDRGMEADIFAVSASEADTNLENKEVNVLLLGPQVRFMKGQFEQKLQPKGIPLD
VINMADYGMMNGEKVLDQAISLMG

Seq ID 195
MKKEIIVYTISDSLGETSQKLLAAASAQYPDISFLNRYNFSFVTTEEELLEILKDALKDKALVVSTLVSKQLITAAKEFSE
RTGLLYLDLMAPFFELIQAKAGVDPIEEPGRRHQLDRAYFDKISAIEFAVKYDDGKNPQGFLDSDILLLGVSRTSKTPVSM
YLANQGYRVSNLPLIPEVPLPPILEEMDPQKMIGLVCSPETLGQIRSSRLASLGLGNETSYTNVERIEQELAYAEEIFAKY
GIPVIDVTAKSVEETAFLIKEKLDERN

Seq ID 196
MEWIEIKHATQNNLKNISVNIPKKQLTVVTGLSGSGKSSLVFDTLAAESRRELNDTFSSFVQNYLPKYGRPEVEKIENLPV
AIVIDQKKVAGNSRSTVGTYTDIYTFLRLLFSRAGSPFVGYSDTFSFNHPDGKCPTCDGLGKITEINLHQLVDYDKSLNKG
PIDFPTFTVGNWRWKRYAHSGLFDLDKKIKDYSPEELALFLYAPQQKLANPPKEWPHTALYEGIVPRMQRSILHTDEGKRH
QKYLNHFVTVKRCPDCLGSRVNERVRSCKINQKSIADAVDMPLTELHSFIRSMDLSLIKTIQEELLVRLEALINIGLSYLT
LGRATETLSGGEAQRIKIAKYVNSALNDIMYILDEPSAGLHPKDIERISRALLNLKNKGNTVVLVEHNPQLIREADFIIDI
GPFAGENGGHVQFSGTYDAFLVSKTLTSQALQEPLPLNDQPRKARKSLSIEHATLHNLNNLSVEVPLGVLTVICGVAGSGK
SSLAEEIYQKAQADNQEIIHLSQKSITANLRSTPMTYLNIFDKVRKLFAEENHVSPALFSYNSKGACPTCKGKGIIVSDMS
FMEDVTSICETCHGTRYKEEVLHYLYNGKNIVEVLALSVKDGYDFFKDQPFALSLKNLLEVGLSYLKLNQSLSTLSGGELQ
RVKLADTLHQKKAIYLMDEPTDGLHLIDIQQSLQLFNRMVEEGNSLILLEHHIDVIKSADWLIELGPEGGKNGGQLLFTGT
PANMLNSTHSITKGYL

Seq ID 197
MKKLKMLGCVGLLLALTACQAGTGNSADSNKAAEQKIAISSEAAISTMEPHTAGDTTSTLVMNQVYEGLYVLGKEDELELG
VAAEEPAISEDETVYTFKIREDAKWSNDDPVTANDFVYAWQQVASPKSGSIHQALFFDVIKNAKEIALEGADVNTLGVKAL
DDKTLEITLERPTPYLKSLLSFPVLFPQNEKYIKEQGDKYATDAEHLIYNGPFKLKEWDNASSDDWTYWDAEKVK
LTEAKVSVIKSPTTAVNLFDSNELDVVNKLSGEFIPGYVDNPAFLSIPQFVTYFLKMNSVRDGKENPALANNNIRKALAQA
FDKESFVKEVLQDQSTATDQVIPPGQTIAPDGTDFTKLAAKKNNYLTYDTAKAKEFWEKGKKEIGLDKIKLEFLTDDTDSA
KKAAEFFQFQL

EENLDGLEVNVTQVPPFTIRVDRDQTRDYDLELSGWGTDYRDPLTVMRIFTSDSTLGGVTFKSDTYDQLIQETRTTHAADQE
ARLNDFAQAQDILVNQETVLAPIYNRSISVLANQKIKDLYWHSFGPTYSLKWAYVN

Seq ID 198
MKQLKKVWYTVSTLLLILPLFTSVLGTTTAFAEENGESAQLVIHKKKMTDLFDPLIQNSGKEMSEFDKYQGLADVTFSIYN
VTNEFYEQRAAGASVDAAKQAVQSLTPGKPVAQGTTDANGNVTVQLPKKQNGKDAVYTIKEEPKEGVVAATNMVVAFPVYE
MIKQTDGSYKYGTEELAVVHIYPKNVVANDGSLKKVGTAENEGLNGAEFVISKSEGSPGTVKYIQGVKDGLYTWTTDKEQA
KRPITGKSYEIGENDFTEAENGTGELTVKNLEVGSYILEEVKAPNNAELIENQTKTPFTIEANNQTPVEKTVKNDTSKVDK
TTPSLDGKDVAIGEKIKYQISVNIPLGIADKEGDANKYVKFNLVDKHDAALTFDNVTSGEYAYALYDGDTVIAPENYQVTE
QANGFTVAVNPAYIPTLTPGGTLKFVYFMHLNEKADPTKGFKNEANVDNGHTDDQTPPTVEVVTGGKRFIKVDGDVTATQA
LAGASFVVRDQNSDTANYLKIDETTKAATWVKTKAEATTFTTTADGLVDITGLKYGTYYLEETVAPDDYVLLTNRIEFVVN
EQSYGTTENLVSPEKVPNKHKGTLPSTGGKGIYVYLGSGAVLLLIAGVYFARRRKENA

Seq ID 199
MKTPTFGKREETLTYQTRYAAYIIVSKPENNTMVLVQAPNGAYFLPGGEIEGTETKEEAIHREVLEELGISVEIGCYLGEA
DEYFYSNHRQTAYYNPGYFYVANTWRQLSEPLERTNTLHWVAPEEAVRLLKRGSHRWAVEKWLAAAS

Seq ID 200
MARVESFELDHNTVKAPYVRLAGTEQNGDALVEKYDLRFLQPNKDALPTGALHTLEHLLAVNMRDELKGIIDISPMGCRTG
FYMIMWDQHSPQEIRDALVNVLNKVINTEVVPAVSAKECGNYKDHSLFAAKEYAKIVLDQGISLDPFERIL

Seq ID 201
MNIPAVRLKEALTAKNIKPSDLAKKTGIGKSSISDWLAGRYEAKQDKVYRIADALDINEAWLMGQEVPMEKNASTIDRIYK
KLEPQRQAIVYQFAEQQLHEQQTQAEILSFPRRDEMTLAAHAGDPEKIFSKEEIEKIHDYLDEIDAKYQQSISSDKKED

Seq ID 202
MYKFVDTNQATHSTPLPSEALNFNGQFLEKVIPGYQTLSVSGRELVPSEIESYQLGIRDGKRHVYARIPERELTVKYRLSA
VNNEAFRDAFNHLNVALFTEKDVSIWFNDEPEMLWFGSKSSVSDVPEGVNQVTGTFTLLLSDPYKYTRSDATSVMWGSPTI
TFQANYLMGNTGSGAFDPPILIEGGAYWGSTMITFQNRAYTMGDLGKEVRPIEIYPTVEGLKVKPTIILTGTGRGVWIKTR
NDTINLGDFDRSEIIIDTENFYLTKNGAPMIRPMNDFYLYPNEPLYIQAKDSDFRLTIRYPNRFV

Seq ID 203
MDFFSSEEIVASLGDKAEMKGKLSFVRLCILGIMAGFCIALGYLAFIRITGTAPTEWGSFNNFLGGALFPVGLIALTFVGG
ELATGNMMVMTLGVLQKKVRVGALCYNWIVVLLTNCLGGFLVAYLFGHIVGLTEGAFLDKTVAVAQAKIADPPLVAFVSGI
GCNIFVCLAVYLGALAKSYLGKMFGLWFPVMVFVVCGFQHVVANAFIIPAAIFSQSTTISWWDYLQNTLWVFLGNAVGGSL
FMAVPLIFMTKPATVKPRVEKTIQTEELYGN

Seq ID 204
MNQRILSTLGFDKVKQQLLQFIVTAQGTNEVSELLPIADENKIQSWLNETQDGLKVQRLRGGIPIPKLENIQPHMKRIEIG
ADLNGIELAQVGRVLSTTSELTRFFDELSENEVDFERLYMWREQLEVLPELNRQLKQAIDDDGYVTDEASPALKAIRQNIR
RSEQTIREELDSIIRGKNARYLSDALVTMRNERYVIPVKQEYKNIFGGVVHDQSASGQTLPIEPKQILEMNNRLRQQQIAE
RNEITRILAELSAELVPYRREITHNAYVIGKLDFINAKARLGKELKAVVPEISQANHVVFKQARHPLLNPEKAVANDIVIG
EEYQAIVITGPNTGGKTITLKTLGLLQLMGQAGLPIPVEEESKMGIFTEVFADIGDEQSIEQSLSTFSSHMTNIVSVLKKV
DHQSLVLFDELGAGTDPQEGAALAIAILDSLGAKGAYVMATTHYPELKVYGYNRAGTINASMEFDVDTLSPTYRLLIGVPG
RSNAFEISKRLGLDNSIIEAAKQIMDGESQDLNEMIEDLENRRKMAETEYLEARHYVDESAALHKELKEAYQVFFEEREKE
LQKARKEANKIIAEAEENAETIISDIRKMQLESGQQGGVKEHQLIDAKTQLSQLHHEETKLAKNKVLKKAKEQKKLKAGDE
VIVNTYGQRGTLLKDNGKGQWQVQLGILKMNVSEEDMTPVAPQKEAKPRVTTVRSAESSHVGTQLDLRGKRYEEALAEVDQ
YIDAAILAGYPQVTIVHGKGTGALRTGITEFLKNHRSVKSYEFAPQNQGGNGATVVKFQ

Seq ID 205
MKEITGATRLAGLFAKPSQHSISPLIHNTAFQNLGVDARYLAFDVGQETLPQAIEAIRTFHMLGANLSMPNKVAAVSYMDE
LSPTAQLVGAINTIVNKDGKLYGDSTDGTGFMWSLKEKKVDVFQNKMTILGTGGAALSIIAQAALDGVKEIAVYNRKSAGF
NDSQKKLANFTERTNCVIHLNDLADTEKLAKDVAESVLLVNATSVGMHPHAHSSPIENYAMIQPKLFVYDAIYNPRETQLL
KEARLRGAETSNGLDMLLYQGAAAFEQWTGQKMPVSVVKRKIENR

Seq ID 206
MTKGYVKSVTELIGQSPVVKLKRMVPEGAADVFVKLEFFNPGGSVKDRIALSMIQQAEADGRLKPGQTIIEPTSGNTGIGL
AMVGAALGYPVKIVMPDTMSIERRKLMQAYGADLLLTPGAEGMKGAIAKATALAEEHGYFMPLQFNNPANPMVHEQKTGKE
IVDVFGKRGLDAFVSGVGTGGTVTGVGHELKRIFPDIEIVAVEPTESPVLEGGEPGPHKIQGIGAGFVPEVLDTTVYQKVA
AVSSEDALETARLMGPKEGILVGISAGAAIKAAIDLAVELGAGKRVLALVPDNGERYLSTALYEFPE

Seq ID 207
VKNQHLTTSQGSPVGDNQNSLTAGEFGPVLIQDVHLLEKLAHFNRERVPERVVHAKGAGAHGIFKVSQSMAQYTKADFLSE
VGKETPLFARFSTVAGELGSSDTLRDPRGFALKFYTDEGNYDLVGNNTPIFFIRDAIKFPDFIHSQKRNPRTHLKSPEAVW
DFWSHSPESLHQVTILMSDRGIPLSFRHMHGFGSHTFKWVNAAGEVFFVKYHFKTNQGIKNLESQLAEEIAGKNPDFHIED
LHNAIENQEFPSWTLSVQIIPYADALTMKETLFDVTKTVSQKEYPLIEVGTMTLNRNPENYFAEVEQVTFSPGNFVPGIEA
SPDKLLQGRLFAYGDAHRHRVGANSHQLPINQAKAPVNNYQKDGNMRFNNGNSEINYEPNSYTETPKEDPTAKISSFEVEG
NVGNYSYNQDHFTQANALYNLLPSEEKENLINNIAASLGQVKNQEIIARQIDLFTRVNPEYGARVAQAIKQQA

Seq ID 208
MAENYQQAAKDIIQLIGMDNIISVTHCQTRLRFILKDHEQVDGKQLEKIDLVKGVFYNGGQYQVILGTGIVTKVYDEIEKL
GINVVSKAEQTAILKNNETGMRKTMRILSEIFIPIVPVIAATGLFLGLKGVIFNDTFLQLFGASVANIPESFQQIVSVITD
TVFAFLPALIVWSTFKAFNATPVIGIVIGLMMVSPILPNAYAVATPDSGVKAIMAFGFIPVVGAQGSVLSAIAAGIIGAKI
ELFFRKKMPNILDQIFTPFMTMLITFLIMILGIGPILETVELGMVDVVQWLIGLPLGLGGFVIGASYPLMVLIGIHHTLTM
VETSLLANTGFNALITICAMYGFANVGSCLAFAKKAQDSKVKSTAIGSMLSQLFGVSEPVLFGLLIRWNLKPLLCVLFTSG
LGGAILAIFHIQSNSYGLAVIPSFLMYIYSAHQLVIYLLVALLSVGVCYALTSLFAIPQEVLISDKVIEEEEREVFEMQHN
TLDEQLFSPVTGYAMNLTAVNDPVFSSEMMGKGLAIMPTANKVYAPADGLLNLVAETGHAYGIQTDAGAEVLIHIGIDTVT
LGQEVFQTQVTQGHRVKKGDLLGTFDRKAIKEAGLDSTVMVIITNTSSYLSVEPMMSDHNEITPEQIILNLNTPN

Seq ID 209
VELVDKDLRSIQETRNLIRKAKEAQQVLATFSQKQIDAIVQAVSEATFNQREKLAKMANEETGFGIYEDKIIKNAFASKVV
YDEMKDKATVGVIHDDAAKKVTEIAVPVGVIAGLIPSTNPTSTVIYKALISLKAANSIVFSPHPNALKSIIETVEIIQKAA
IAAGAPEGCVSVIKTPTMQATSELMKNKETNLILATGGNAMVKAAYSSGTPAIGVGPGNGPAYIERSANVPHAVKQIMDSK
TFDNGTICASEQSIIVETVNREAVKEELIKQGAYFLSPAEADKLAKFILRPNGTMNPQIVGRSVQHIASLVGLSIPKDRRL
IVAEETHVGLKYPFSREKLAPIIAFYTVENWEAACALSIEILKGEGAGHTMGIHTENKEVIREFGLRKPVSRLLVNTSGTL
GGIGASTNLVPALTLGCGAVGGSSTSDNIGVENLFNLRRVAYGVRDLEEIRQEFGQTSTTSVATSCETTNQEELVNAVVAQ
VLARLN

Seq ID 210
MNELNKTPKEIVKELDQYIVGQQAAKKSVAVALRNRYRRLQLEENMQQDITPKNLLMIGPTGVGKTEIARRLAKIVNAPFV
KVEATKFTEVGYVGRDVESMVRDLVENAIQIVEKQQYSRVYAQALKKANQRLVKVLVPGIKKEQKQAGGNQFEQMMNMFNM
AQQQQEAQEEVTEDIRTNRRTILEQLEKGLLDNREVTIEIEEPKKTMPAMNNGLEQMGIDLNETLGALSPKKKIERTVTVR
EAQELLVKEESAKIVNDADIHSEAIRLAESSGIIFIDEIDKITSKSQQNSGEVSREGVQRDILPIVEGSQVNTKYGPLQTD
HILFIASGAFHLSKPSDLIPELQGRFPIRVELDDLTADDFVSILTEPNNALIKQYVALIGTENVSVIFTKEAIERLAHIAY
DVNRDTDNIGARRLHTILERLLEDLLYEAPDMQMGEITITEAYVNEKLNDIVQNEDLSRYIL

Seq ID 211
MAKRIIIMNFDIESKSYQAFSEIKKMQAERQL
KGEQMAVVTHVNDGQHQFKINDFIDFTGNNHTSKDSMIGMLVGILGGPLGILFGWFAGSMYGASKDAKEIQEAQTVFEHVI
QKIDEGQTGLLLIAEEEDNRPLNQLVMFDLGGEITRLDLEEVQQEINDANEVANEAKQSWQAKKEQHKEATSKEE

Seq ID 212
MEKQTITIYDVAREANVSMATVSRVVNGNPNVKPATRKKVLEVIDRLDYRPNAVARGLASKKTTTVGVIIPDVSNAFFASL
ARGIDDVATMYKYNIILANSDGDDQKEVTVLNNLLAKQVDGIIFMGHRITDDIRGEFSRSKTPVVLAGSIDPDEQVGSVNI
DYTEATKDATATLAKNGNKKIAFVSGALIDPINGQNRMKGYKEALAENGLSYNEGLVFESEYKFKAGINLAERVRNSGATA
AFVTDDELAIGLLDGMLDAGVKVPEEFEIITSNNSLLTEVSRPRLSSITQPLYDIGAVSMRLLTKLMNKEEIEEKTVVLPY
GIDQKGSTK

Seq ID 213
LKKYNVDWNYWVVRFLFVMALIVGYLLITNYQHFVHSVSGLLGILSPFITGFVIAYLLSGSQKKIEGLLERVPLPVVKKAK
HGLSVLLLYLIILFIFVLTLNYIVPLLISNLVDLANSLPTFYDHMVQFVMSLEDKGILKTAAIEKYLNSVLKDLSPERFLN
QWTQALFSLGTLTKNVSSFFLNAFLTLIISIYALVFKQSILTFVEKAAHKLLSEKVYKQTQTWLNTTNKIFYKFISCQFLD
ACIIGVSSTILLSILNVKFAVTLGILLGICNMIPYFGSIFASIVAGVITLFTGGVTQAITVLLVLLILQQIDGNIIGPRIM
GDALNVNPILIIVSITIGGAYFGVLGMFLAVPVAAIIKIIVSEWLNESKENDKIVDSIES

Seq ID 214
MKKFSIRKISAGFLFLILVTLIAGFSLSANAEEYIVPAESHSRQKRSLLDPEDRRQEVADTTEAPFASIGRIISPASKPGY
ISLGTGFVVGTNTIVTNNHVAESFKNAKVLNPNAKDDAWFYPGRDGSATPFGKFKVIDVAFSPNADIAVVTVGKQNDRPDG
PELGREILTPFVLKKFESSDTHVTISGYPGEKNHTQWSHENDLFTSNFTDLENPLLFYDIDTTGGQSGSPIYNDQVEVVGVH
SNGGIKQTGNHGQRLNEVNYNFIVNRVNEEENKRLSAVPAA

Seq ID 215
MKRKKIKNQLLVLTSILVLLSLEVAPVVTLAAELPSSSLQTALSSETTITSEEKVTETTETTVATSTTSTSSSSSESSSST
DTTTESTSQSTTETTTTNTSSETKKEPTEPTVSSEITQVEQSQPPQVPVTKQEPEEPIQVPEANNNFVEENQAVSLNPSL
KVDEIASSNLKGYELPLLSSFGEKKRAVVVAEALRHVGKTKKEFNLTEQALTSSFLAQKIYQQLFKIDIGSTPQEQMTFGK
VRSIEEAEPGDLIFWQTAEGKTLQNGVYLGQGKYLIAAAESDSKEKPEVIAQLENIYTAKQQPDSKEEKRLVVTNPFKEPI
LTEYGKEVLATYGASFEMQKSEQTTAFIKKIGETARELGEKYDVFASVMIAQAILESGSGESQLAKEPYYNLFGVKGSFQG
NSVSFSTKEADQRGQLYTISAGFRDYGGYNDSLQDYVQLLRQGIDGNQDFYKPAWRSEAKNYLQATRFLTGKYATDKQYDN
KLNSLIAVYNLTQFDLPKTVDGLIIQSKNKLSEAEQQQMHFPVYDGINYNRSGSYPVGQCTWYVYNRFKQLGTSVDEFMGN
GSDWGRKGRALGYQVSSLPKAGRAISFQPGVAGADNQYGHVAFVEAVTSDGIIISESNVINDQTISYRVLPNVIAYSSGVT
YIGA

Seq ID 216
VDAEAYLGPDDEAAHSFGLRKTPRLQAMYDKPGTIYLYTMHTHLILNMVTQEQGKPQGVMIRAIEPVEGVDKMIENRQGRQ
GVELTNGPGKLVAALGIDKQLYGQSIFSSSLRLVPEKRKFPKKIEALPRIGIPNKGRWTELPLRYVVAGNPYISKQKRTAV
DQIDFGWKDEENEKSNNAHILRGTT

Seq ID 217
MAKKKNVSVISVEKPTWFPLTDETGAFPVTGAPITIGTAVSIKPDVTTETTPDYGDSVVQDQYVAFGGAEVTLETNGYQNE
VLAEITGGKKLKGGVLRSADDIASDGAFAYRRRKSNGKYRYTIFYKGKFALTSDETSTLEGSSVSYTHPEWTGSFVDVPGL

GYMYSVDEDDEGVDLEMIKNWFTEVMDPRKENTTAVTGVTLDQTELNLKVGQTATLTPTITPDNASNKKYQFRSESEAIGT
VTPIQGKVTAVGEGTTEIVVTTEDGNFTAKCTLNVTTAD

Seq ID 218

MFLGLIFILFGVFGLFRLGFLGTLLANCLRLVVGNTFPFAAILLILYGLLVMIYGKDFPLKRGRPIFGAVLIYISVLLFFH
AFMFRNVSGSQPDILGNTWEFLQSDLKANQVTQNVGGGMIGALLYQGTYFLVAQFGSYLIATLLLLAGIFLMSMWDFQQIV
DHFQSIQDRLSHVSAKSQARQEEKEAKRAAKKEAKAAERQAKIEAAAQQKLQERERMEQAAAERLTKTPVETHQPMVEEPA
APTPVQIDSFQQQNQAMFVPPIAATKPQREQEEEAADEAGVLEFEISEEAERDRDYQLPPTDLLDTIQATDQSGEYEKIEKN
IGVLEQTFKSFGVDAKVVKASLGPSVTKFEVQPAVGVKVSKIVNLTDDIALALAAKDVRMEAPIPGKSLIGIEVPNSAIST
VSFRDIVEAQPSHPDKLLEVPLGRDISGMVQTADLSKMPHLLIAGSTGSGKSVAINGIITGILMQAKPHEVKLMMIDPKMV
ELNVYNGIPHLLTPVVTNPRKAAQALQKVVQEMEFRYEKFAATGVRNITGYNQLIQQKNAEDGENRPILPFIVVIVDELAD
LMMVASNEVEDAIIRLAQMARAAGIHMILATQRPSVDVITGIIKANVPSRMAFAVSSGTDSRTIIDTNGAEKLLGRGDMLF
LPMGENKPIRIQGAFISDQEVERVVAFVTDQQEAEYQESMMPTDEPTTSGGGEAPQDELFEEAKNLVVEMQTASISLLQRR
FRIGYNRAARLVDELEAHGVIGPSEGSKPRKVFLQAESEEAATETPEQ

Seq ID 219

MRKSFNLAVQALTVQYQGRTALNNIHVTIPSGKITGIIGPNGAGKSTFIKGLLGLIKTKERDVLLNNQAIDQQKTTIAYVE
QRSALDLSFPISVFETVLLGTYPNLGLLKRPGKKEKQAAMAALKMVQLEDYAQRQIGELSGGQLQRVFIARVLAQGAEVIF
LDEPFVGIDMSSEKVIMDILKSLKNQGKMIIIVHHDLHKVSHYFDELIVLKNRLIAAGPVEQTFTAETLQEAYGDLLGDLL
IQGVAK

Seq ID 220

MIGLELTFTVPQADEVIRSLVETYHDEKDIVIGAGTVLDAVTARLAIMAGAEFIVSPSFNRETAEICNLYQVPYLPGCMTI
TEMQTALKSGVDIVKLFPGSTYGPSIISAFQAPLPYLNIMPTGGVSLDNMASWFEAGVTAVGVGGNLLAHAKTGEFGKVTA
LAKQYMAKFREIKEN

Seq ID 221

LAGVTKEQIAKAKEWDLLSYLQAYEPQELKKSGPREYCTRTHDSLKISNGKWCWNSRGIGGRTALDYLIKVRGMDFVGAVE
TLCGYCAPPPEKAPPIKEKQPKPFRLPQASRCASAVVGYLQDRGIDPELLGVCMEAGILYESCKYQNCVFVGRDTAGNARA
ASLRGTRDGFRMDVEGSDKRYSFCLPAGKADCPRLAVAESPIDALSLATLVKLSGGEPRDSHYLSLGGTGPRALIQFLHDH
PHVTQVSLCLDNDKAGLEGMERLTQEIQNDAELSRRVKLVYPNPPKQGKDYNEFLCTHVKAVRTAQRQRDAAR

Seq ID 222

VRIKEGRTVGLFTKVSPEEKAAIDRKMELLGTSNLRAYLRKMAVDGYIVQLDMGAVLELVKLLRSVSSNVNQIARRCNSTH
NLYEQDVEDLRQGYSEIWQGVNDLLKKFEAL

Seq ID 223

MANKLQTVSQLADQTAKAVTRNAEGWKSYLTTASRLYKYSFDEQLLIYAQRPDATACASMELWNEDMRRWVKAGSKGIALI
KKDGIRPRLTYVFDVADTRPVRGAKMPYLWEMKDGHHAAVMDTLARRYGETEKKDIGSALMEQAKHAVEEVYREHLSDLAY
DAQDSLLEGLDDFNLEVRFRNLLTASVQYTLLSRCGLNPADYLEDEDFDGIHEFSTPAVLHHLGSAASEVSMNMLLEVKKA
IRQAEKEKAQNRQKIPKNPEKPLAKEPVIGYTKVKEQFNTLKRESEERSIQNGRTGIQEDGRLPDSRLGDGRGGRDGGNAA
GQVRQAAADLSSGTPQGDIHLDAADRAAGTPPAGDRPAGAGTGRPDRGGIKETERRGRGDESPRPDGMGAGSQPVSRPGGG
NRTERDRLQVNQENQQEAAGEQSAVSASEKPAFTQFSLFPTVEQQIETIAQAQKTEQELRPAVSSGKVTDAVIGRALTSGG
NEPDSILRIVAFFQKDPTEQSAADFLQKEYGSSGGKGLKIAGHEYSLWFSHSGIHIAPGRSAGGGSLVTWKQAAAQIRQLLN
SGQFASQDKIDAARDNEFRELSEKLWFLRQNFSEEAKEKGLLPTIDALYGGFPDSTAKIAALLKDPAERAKIAQEVRQFAI
SHQENPNLLRFRQRPAPSELFTRLTDMDKPVTAFHGVEGFAPARGAFVTEDEITRMFTGGSGVSEGKFRIYAYFMQGHDAK
ECIEFLKNEYGIGGHGYIGYDEWHDGKGIKLSRADDFSGGNYDTVTLNWKQVQKRIAGLIKTDRYLNRREKAYLPEYEKMQ
LARSLYTFQYYDPNDASKTIPHDWDVDAAKKVFRPLLDDPEQCAAHYEKMVKALAMVSPDERAYSLMEPTLQKMGAYLRGE
YSLFTPLPDAVLQEERQKKQKQKQKERQDSARQTAEPATGLAAAAKALSRKKQPAPREDADGQLTLDMFGLSSEPEQPAAT
PEPEPELSQAETGASETAVPIASKPEPPQETPNIPEGAAPIMAESPADRYDLGYGHMGNGLTVWNRLEEEHGDYKTVAHIA
PDRTVTFYDADMPEEIREKIQKVAAATEMSISATQDTPVFSTPPQEPERVQPGNSEPEPEKVQDAPAAVNPEPETGDSNIV
PSPAQKAGAEPTGTGSPQTQEAPKTAAADGLNLTPNVEEYLNLKAQYPDKLVGVRVGKTMLFYGTDAEEAAPALGTKTITR
DIPDLGMVSITGANGWQSVLKKLLEHGKSVVLARPDTERGGDAPYEIIKERSAADYIPVGMELTIDGRLMKIDSVDYNAGT
VSLQDMELRGWYPIFRSESIPFVRQFVEEVQQEHFTAEPMQEPERPETSDLDTAKQLIEQFAYAEYGSDDVDFSDLEHIGI
AYGTTEDGGLEVQVDVNLLDFSISQSVDGKCVETRQYGSLRELIDMELAFLDYDMLVSVEPDIEERLKAELNQRIRWSEMA
GAREGVEPTEPEIFFPEKDVSPEQEAPPELVEIDGGQITETPAQRQTRRRAQEEVDSRVFPSEIILQPLRLEPERHNFRITD
ENLGAGGEKTKYQYNVEAIRTLKQIEAENRLATPEEQAILSRYVGWGGISHAFEPNDPKWAKEYAELKELLTPGEYQSAQS
TVLNAHYTSPTVIQAIYNAVEQMDFTPGTVLEPSMGIGNFFGMLPEKLAAAKLYGVELDDLTGRIARQLYQKADITVDGFE
RTDHPDDFFDLAVGNVPFGSYQVHDKRYDRQNLMIHDYFITKTLDKVRPGGIVAFITTKGTMDKENSKAREALAQKADLLG
AVRLPSNAFKANAGTEVTTDILFFQKRDRIPEKLPEWVEVGQTEDGIPLNRYFLDHPEMVLGTMTMGRSMYGNETETACQP
IPGADLSGQLAEAIRHIAPPDRELLEVDSGQDGEELESIPADPTVRNFSYTLSNDKLYFRENSRMTQAVLGKTPTERVRGM
IGIRDSARRLIDLQLAGADDTEIQSEQAKLNRLYDAFSAKYGLLNSTGNKLAFEQDSSYPLLCSLEIINEEGKLERKADMF
TRRTIQTHRAVTSVDTAVEALAVSIGEKACVDLGYMASLMGGPEKIPQIVEDLKGIIFKNPATGPFDLEDGGAHWHQGWQT
ADEYLSGNVRAKLAIARAAAEENPQFAINAEKLEQVQPKDLTASEISVRIGASWIDPEYYQQFMFELLHTPVYLQDRKIKL
QYSPTTGEWNVQGKSTDNRDNVRVYATYGTKRINAYEIFEQTLNQRDVRIFDKTEVDGKEVRVLNEKQTAIAQQKQRAMCE
TFKDWIFKDPERRETLCRRYNEKFNCIRPREYDGSHIRFAGMNPEISLRTHQENAVARMLYGKNSLLAHCVGAGKTPEMIA
AAMEGKRLGLNQKSLFVVPNHLTEQWGGDFLRLYPGAKVLVATKRDFEPARRKKFCARIATGDYDAIIGHSQFEKIPLSP
ERQRKAVIKGQIDEIVTAIAEAKAEDGERYTIKQMEKTKKNLEAKLQKLADGKKKDSVVTFEELGVDRLFVDEAHGFKNLFL
HTKMRNVAGIAQTDAQKSSDMFAKCRYMDEITGGRGIVFATGTPVSNSMVELYTMMRYLQFDTLEQNGHRHFDAWAADFGE
KVTAMELKPEGSGFRSKTRFAKFYNLPELISIWKEAADIQTADMLNLPVPKAEYITVTTEPSGFQKEMVEELGERAESVRG
GQVDPHIDNMLRITSDGRKLALDQRLQNPLLPDDPDSKVNACVKNIVQEWQASTEILGTQLVFCDLSTPKGDGSFNVYDDI

KEKLIAKGIPEAEIAFIHDANTETQKAELFAKVRRGQVRVLIGSTAKMGAGTNVQNRIVASHDLDCPWRPADLEQRAGRSL
RQGNMNASVKMFKYVTKGTFDAYNWGLVENKQKFIGQIMTSKSPARSAEDVDATALSYAEVKALATGDDRIREKMDLDVQV
SKLKMLKANHTSQQYEMQDKALKYYPQKIAETKLLIEALTADLPTVQAHPVKDDAFSMTVMGQTYTERKAAGEAVIKACML
MDDPEKTVDLGEFRGFPMQLRCDGSKFRVTMKQNLTYSAELSDDAVGNVTRINNALESLTERLDAQKARLATLESELENAK
EEADRPFPKEEELREKSARLNQLNRELSRPDKKEDEQEVEESPDLEDAGKPAPVVPMPAVAAGYKPSIRQAIRNYEPPTPV
HSGMERTPRKEAVL

Seq ID 224
MPVLTSGQQERKSKQRKIRNSEYYDMEGTFDRLYAESKKDKTFNHLMEIIESEENIKLAYRTIKKNTGSDTSGVDKRTIAD
LAKLSEEEYVRLIRKQFSNYHPGPVRRVEIPKPNGKTRPLGIPTIVDRIVQQCILQVMEPICEAKFSENSNGFRPNRSAET
AIAQCMRLIQVQHLYHVVDLDIKGFFDNISHTKLIRQIWALGIRDKKLLCIIKEMLKAPVVLPNGEKTYPARGTPQGGILS
PLLANIVLNELDWWIASQWEEMPTKTKFKTRSNAQGTEIKSHAYRALRRSRLKEMHAVRYADDFKIFCATHEDAVRAYKAT
ELWLKDRLGLEISPDKSKVVNLKRQYSDFLGFKLKVRKKGKKYVVRSHMSDKAYKKAHEKVSEEVKKLAYSSDDNAQFMQL
QKYNSVVAGLHEYYCISTEVSHDFSRLAFSINKQLRNRLKGDLSKKGQLRNGFIKEKYGASRQMRFLHGRPVVPLGYVQSK
NAQHKRKSINKYTVKGREQIHKNLAIDTATMLWLMRNPVKGRTIEYADNRISLYAAQYGKCAVTGIPMNSHDIHCHHKVPV
SNGGGSDEYANLILVSKAVHILIHASSEPTIEKYLKSLNLDNKQIEKLNKLRSMAEMPPIIL

Seq ID 225
VTEKARDLFDKLMTQIQLDDTAKEHPLIQNGRIDKVIVHQQSRLWEFHLSFDELMPVMLYQTFMQQLELAFQQIAQVSVQI
TTNQTTFTEEQLTDYWQLALLNSQCNTPLVQKVLKTQTPQFEDRKIILPVDNEAVIPYLKQQVLPIIEELYFSYGFPKFHI
EPKMDQQQAAEVLKKFEEQKLEQAAAFQQQAAESLVKHEQMKKEKQQQAPAFDGPIRLGRNIPNDEPIMPMGNILEEERRI
TIEGFIFDKEVRELRSKRKILILKITDYTSSFVVKKFSNGEKDEQVFDAIQPQSWIRVRGSVQEDTFMRDLVMNAQDLMEV
AHAPRKDYAPEGEKRVELHMHSNMSTMDATNKVGDLVAQAGKWGHKAIAITDHGGAQAFPDAHAAGKKAGVKILYGVEANI
VDDGVPIAYNEEHIELTDATYVVFDVETTGLSAVYDTIIELAAVKMHKGNVIETFEQFIDPGHPLSQTTINLTGITDEMVR
GSKSEEEVLRMFKEFSEGTILVAHNASFDMGFLNTSYGKHGIPEAANPVIDTLELSRFLYPHFKSERLNTLSKKFGVNLEQ
HERAIYDSESTGHLCWIFLKEAKENHDMHFHDDLNRHIGEGDSYKRARPFHATILATTQAGLKNLFKLISMSNVDYFFRVP
RIPRSQLSKLREGLLIGSACSNGEIFEAMMQKGVEEAKNRAKFYDYIEVMPKPVYAPLIEQELVKNEADLEEIISNLVKIG
DELGKLVVATGNVHYLNEEDAIYRKILVGSMGGANPLNRHSLPKVHFRTTDEMLTEFQFLGQDIAKRIVVENPNKVADLCE
EVIPVKDDLYTPKIPGSEQEITDLSYNRARELYGDPLPEIVEKRLEKELNSINGNGFSVIYLISQKLVHKSNEDGYLVGSR
GSVGSSFVATMTGITEVNPLAPHYYCPECQYSEFYEDGSYGSGFDMPEKACPKCGARLFKDGHDIPFETFLGFHGDKVPDI
DLNFSGDYQAEAHNYTKVLFGEEYVYRAGTIGTVADKTAYGFVKGYERDHNLHLRGAEIDRLAKGSTGVKRTTGQHPGGII
VIPDYMDVYDFTPIQYPADDQEAEWKTTHFDFHSIHDNILKLDILGHDDPTVIRMLQDLSGIEPKTIPTDDPEVMRIFEGP
DVLGVDAEQIYSKTGTLGIPEFGTRFVRGMLEQTHPSTFAELLQISGLSHGTDVWLGNAEELIRRGEANLAEVIGCRDDIM
VYLIHAGLDSGMAFKIMETVRKGQWNKIPDELRDTYLNAMKENNVPDWYIDSCSKIKYMFPKAHAAAYVLMALRVAYFKVY
FPPILYYCAFFSVRADDFDLVAMSQGKEAVKARMKEITDKGMDASTKEKNLLTVLELCNEMLERGYKFGMIDLYKSDAENFV
IEGDTLIAPFRAVPSLGLNVAKAIVAAREEQPFLSKEDLATRGKVSKTLIEYMNENGVLKDLPDENQLSLFDML

Seq ID 226
MERSNRNKKSSKKPLILGVSALVLIAAAGGGYYAYSQWQAKQELAEAKKTATTFLNVLSKQEFDKLPSVVQEASLKKNGYD
TKSVVEKYQAIYSGIQAEGVKASDVQVKKAKDNQYTFTYKLSMSTPLGEMKDLSYQSSIAKKGDTYQIAWKPSLIFPDMSG
NDKISIQVDNAKRGEIVDRNGSGLAINKVFDEVGVVPGKLGSGAEKTANIKAFSDKFGVSVDEINQKLSQGWVQADSFVPI
TVASEPVTELPTGAATKDTESRYYPLGEAAAQLIGYTGTITAEDIEKNPELSSTGVIGKTGLERAFDKELRGQDGGSLVIL
DDKENVKKALQTKEKKDGQTIKLTIDSGVQQQAPAIFDKRPGSAVITDPQKGDLLATVSSPSYDPNKMANGISQKEYDAYN
NNKDLPFTARFATGYAPGSTFKTITGAIGLDAGTLKPDEELEINGLKWQKDKSWGGYFATRVKEASPVNLRTALVNSDNIY
FAQQTLRMGEDKFRAGLNKFIFGEELDLPIAMTPAQISNEDKFNSEILLADTGYGQGQLLISPIQQATMYSVFQNNGTLVY
PKLVLDKETKKKDNVISANAANTIATDLLGSVEDPSGYVYNMYNPNFSLAAKTGTAEIKDKQDTDGKENSFLLTLDRSNNK
FLTMIMVENSGENGSATDISKPLIDYLEATIK

Seq ID 227
MKKVLMGVLSLGLLLGAATGCTSDQEKAAGKTKASSEKTEATSGASANGYTDPSELKDSYDVVIVGSGGAGMTAALQAKEA
GMNPVILEKMPVAGGNTIKSSSGMNASQTKFQEKEGIKDSNDKFFEETLKGGKGTNDQELLRYFVDHSAEAIDWLDTKGIT
LSNLTITGGMSEKRTHRPADGSAIGGYLVDGLVRNVREEKIPLFVDADVTDLVEENGQIDGVKVKMKDDKEKTVKAKAVVV
TTGGFGANEKLITQYKPELKNYVTTNQEGTTGDGIQMIQKVGGALVDMKEIQIHPTVQQSDAFLIGEAVRGEGAILASQKG
ERFVNELDTRDKVSAAINALPEKSAYLVFDQGVRDRAKAIDFYDQKGFVEKGETIEELAEKIGMPADTLKATIDTWNQDVN
AKDDKQFGRTTGMEADLSTAPYYAIKIAPGIHHTMGGVKINTKTEVLREDGTPIKGLYAAGELTGGLHGQNRIGGNAIADI
IIYGRQAGTQSAEFASAQK

Seq ID 228
VIVRGGGDLATGVIQKLWHVGFKILVLETECPLAIRRTVSVCDAIFQKEQRVEDLVAVRITSLKDCAKCWQQNKLFVFVDQ
TASAIQQLKPLIVIDAILAKKNLGTHRGMAPITIALGPGFSAPQDVDVVIETMRGHRLGRLYFEGTALFNTGIPGEIGGKS
AERVVHAPASGQVTHLKNIGDLVLKGEALFLIDQVPVYSPLTGTLRGLISEKVTCYQGLKCADVDPRPVEKVDCLTISDKA
RALGGAVLEAIFMIGRRKNVL

Seq ID 229
MYSIIVNGQSETCETNKKLMDFLREDLGLTGTKDGCNQGSCGACTVLVNGKASKACLFTLEKLAGKEVTTIEGLSQRQKDV
YAYAFAKTGAVQCGYCIPGMVISAQGLLNKKPEPTKEDIQKAIRGNICRCTGYVKIIEAIQLAAKMFCEEASIPEEHSNGK
LGEDPQRVDAVEKTLGTGIYVDDIDIEGMLHASALRSAYPRAKVLSIDSTQALAHPDCVAVFTAKDVPGNNKIGHLEFISD
WDVMIPEGEITRYVGDAVALVVSKRKETLPEIKNLVEVDYEEMIPLTSCEAALAEGAPAIHEKGNILSHEHLVRGNAEEVL
ENSAFVVTEHYSVPINEHAFMEPECAIAQPEGEGILLFSAGQSIYDEQREVARMLGLDKEKIHVQSKLVGGGFGGKEDMSV
QHHASLAAWLLKKPVKVLLSREESLMVHPKRHGMEMDFTTGCDEEGNLTAMKAVIYADTGAYASLGGPVLQRACTHAAGPY
KYQTIDVEGFAVYTNNPPAGAFRGFGVCQTAFAIESNLNLLAEKVGLSPWEIRFKNAVAPGDTLPNGQLVSKNAALKEALL

AVKDVYEQAEVAGISSFFKNSGVGVGLPDTGRCIISVEEGKIHVRTSAACIGQGMATVTTQIACETLNLPPEMIVAEAPDT
RRTPNSGTTTASRQSLFTGEATRRAAMQLRYELDMGRALSDLEGEEFYGEYSAKTDPLINDKKSPVSHAGYGYAAEVAILD
DKGKVAKFVAAYDMGQVVNPRAAEGQIEGGIAMGMGYALTEKFALEEGYVKAKYATLGLVNATQVPPIETILVHADNLHEG
LAYGIKGVGELATIPTAPALAGAYFALDGQLRPSLPLENTPYQKKKR

Seq ID 230
MSDIMHPISIEALLNWIFSEYQQDGTIFGIRKFYHADPTKTISLFGEKMETPCGPAAGPHTQLAQNIIAAYLTGSRFFEVK
TVQILDGEDLPVSKPCIAAADECYNVEWSTELRVPQAYDEYVKANFVLKLLSKEFELGDPNGFIFNMSVGYDLAGIQSPKI
DRYINKMQNAEGTPIWAECQAAAKKYLSYFKKVDDLYIEAISPKVCHSITLSTLHGCPSDEIERIAAYLLSEKGLHSFIKC
NPTMLGYEYARQTMDELGFDYMVFDDHHFKEDLQFEEAVPMLQRLQLLANSKNLSFGVKITNTFPVTIAANELPGDEMYMS
GRSLFPLSISILAQKLSEAFDGKLQISYSGGADIFNSKEIFDAGIWPITMATTLLKPGGYQRMNQVANVLSAAEYPQMVHVN
LDKLAQVVEKAKTQARYQKSIKLPESTKLRKTVPLTDCYIAPCRSDGGCPINQDIPAYLRYVSEGNYLKALQVIVDKNPLP
FITGTICAHPCMTKCTRQFYEESIHIREVKLEAAEHAYDELLTTLEKPQPKENAPKTAVVGGGPAGISAGYLLAREGMPVT
VFEKSETIGGVCSQIVPEFRISMESVQKDVQLAEFMGAEFRTGQEAPSLAELKNQGYTNVIYAIGAWKHGVLRLESGRALN
SLEFLKANRENPTINPYGEQIVVVGGGNTAMDCARAATTLPGVKKVSVVYRRNKRNMPADEEELYLALEDGVDFLELLSPI
KHENQQLTCEKMVLGERDASGRRRPIGTGEMIDIPADTVIAAVGEKVDTEFYQALGIHTDNYGKVVSNQETLETNIPGVYV
IGDANLGPATIVEAIADATKAANNICLVHNHHYEKDNLNSDVAFVRNKRGILVTDEMSCSQASRCLECSTICESCMDVCPN
RANIVVYVEGKPQIVHVDRMCNECGNCETFCPYASAPYKDKFTLFNSEADFYDSTNSGFYVQNPVDKICLLRLWGTVSTIQ
LTDQGLDVPEDLIALMVTMIEEYAYCMQA

Seq ID 231
MTKQTIKEQILHFMESQKKKSFSMEEIAQGLNLEKSSDFKILVQTIAQMEREKSVSFNKKGKVLLPMKDLLIEGTFRANER
GFGFVTIDPEEPDVYIPKEATNFAMDGDTVLIDVIQHADPFSDRGAEGKVKEIKERAVSQVVGEFVAYSEEEMAEMGLYGY
MIPKDKKLNQYTVSIAPEGIKPVDGSIVIAEITYYPDQEYPTSMEGLVKQVIGHKNDPGMDILSIVVAHGIPTAFPDEVLA
EADQVPETIAESDLVGRRDLRDQLIVTIDGEDAKDLDDAVTVQKLANGNFFLGVHIADVSYYVTEGSQLDMEAYERGTSVY
LTDRVVPMIPQRLSNGICSLNPHVPRLTMSCEMEITPEGEVISHEIFQSVIQTTERMTYTAVNEILEEQKPETLERYKELV
PMFKEMGELHHILEEMRMRRGAISFEDREAKVLVDENGHPKDILLRTRGVGERLIESFMLAANETVARHYHDLKLPFIYRI
HEQPKEEKMQRFFDFAAVLGILVKGTKENISPKDLQKVLEQVENKPEEVVINTMLLRSMQQAKYSEDNYGHYGLAAEYYTH
FTSPIRRYPDLIVHRLIRSYSQDQSEKNQEKWNEALPEIANHSSSMERRAVDAEREVDAMKKAEFMVDKVGETYDGIISSV
TKFGIFVELPNTIEGLIHVNNLKQDYFHFIENHMALVGERTGMTLKIGQKVQIRVEKADPETREVDFELISAEEVAPVEGP
KGRKKGKANSSTRSNNQRRNKKDESFDGKKKKNKKKGKGKKQPFYKEAMKQKNKKGKKKK

Seq ID 232
MKNRYFILMILSFYFFAFGIETSAAELRFSVETEIPENQIDKTKTYFDLMMKPDQEQILKVRAANSTDENLVIDVSVKSAT
TNSNGVIEYGESLTALDKSAPADLSEIIQLKDGGESVELPAKSEKAVELRVKMPKEEFSGQLAGGITFSEKVDETKDKQKE
NTNGLAIEKNRYAYTVAVLLRENETVVQPELSLEKVEPTQRNARSVISATLLNHEAAYLQSMKVTANVKNKKTNNVILEKEQ
EDMQMAPNSIFNFPIPYEENEMEAGTYVLAMIVEGSGKKWQFTKEFTISKEEAKTFNEKDVTVKKTESKLIYLLIGLLILL
LIICLFIILRLKKQKNK·

Seq ID 233
MSRKRKISLISLVIILVFVTVGSAYFAVAGSYLKKTIDKGYVPIKNDYNEAQNKDSQSFLIMGLDNTIERKLGTTRTDAMM
VITVNNKTKKITYLSLPRDSFVQIDAKNYQGMQRIEAAYTYDGPTASVNTVEKLLNIPINHYVVFNFLSFIKLIDAVGGID
VNVKQAFDGVTKDGPGSIHFDAGKQHLDGTKALSYARERHSDNDIMRGFRQQEIIQAVEDKLKSGQSIMKIMDIIDSLNGN
IQTDVDSNELTHLVKEGLTWTNYDKQQLSFDWRTFSNEGRSMVELYPDSIENVRHQLRVSLNLEKPDERDQDGYVFHTNGE
FLYQSDYTVQDEAAEENEMTSINGNTYIGVPGNTQTGPLPSVKTENGFIK

Seq ID 234
MKKSQAVQLIQELSNANGVSGPETEVVRILQHATADFTIQRLDSIKNLYLEKKNNLSEGPVVLFDAHSDEVGFMIQAIKEN
GLLRFFPLGGWVPNTISAQKVRIRNREGTYLPGVVTSRPPHFMTPEERQRPLTIADLTIDIGATSKEEVIETYKIDLGAPV
IPDVTCCYNEQTDLFLGKAFDCRIGCACLVDVMEELKEETLPFKLVATVTAQEEVGERGALIAAKQVNPDLAIVFEGCPAD
DTAETPEMIQSAMGKGPMLRYFDVSMITNPEFQEYALEIAKIHKIPVQVSVRSGGGTNGMAITQVQGAPTIVVGIPVRYAH
TPHCYVDFQDYQAAKELVIQLIKNLDADKIQALVQPLSKEWNK :

Seq ID 235
MTTSQMIADFTTLAIQAGGWMELDRLYLQNRLLSMIGEQELGEVDIRPVATPAADLAEQLCQVASANQLVKTEQQKEQFMV
QLMDLLTPPPSVVNAFFAQHYAKEPQEATEYFYQLCQKNGTVIEQEEPVVFSTVYGDFLANKVHSEASKATLSAQSYPRCE
WCMATEGYQGSQQFPATTNHRVIRMNLDGESWGFSFVKQAQYQQQGVIAFEKHQSAKRSIKTFQQLLKIVEVFPHYFAGID
ADFEQNEHVYYQTGLQQFPLAEASISEYVELANYPLINAGMVNWPVETFRLEGPNASEVAQAANDIFEQWQMLKLPTDEIQ
IVARRKELLYVMDLIFSRPQAKPSLTLAEVQGLTTWNNQKTQALETVASAYQQRLKEASAFAETSEGKAAFLAMVAPVTH

Seq ID 236
MSILWVINGILLVIVLAMVFNELYLKIMVKRSAKMLTEEEFKETMRKAQVIDVREKDTFDAGHILGARSMPYSMLKTTIGS
LRKDQPVYLYDQKKALSIRAANLLRKNGYTDIYILKGGYDGWTGKVKKRNS

Seq ID 237
VIFMNPDEPNFIWKDLNAVRDMGCIIENELSEVLPNKRYETYSIIGRSGEFNETFNDYEPFDYEIEDVTIPYENLAAVKRW
LTGKSKLITHNDEDKYLDAICTMSKPTSFKNEWGVFYTFNVEFRCQPFKRKVNEQPKVIKTKSIEITDHGDEIAFFYIEIN
SKGGDITLNIGSNSLTILRTQSGIVTIDTEKGKAIQEGNPLFTRGSWIKTNPGQNKLNISGNFIEAKFWNRSAYL

Seq ID 238
MSIHITFPDGAVKPFDSGITTFDVAKSISNSLAKKALAGKFNGVLIDLDRPIVEDGSLEIVTPDHEDALGILRHSSAHLMA
NALRRLFPNIKFGVGPAIDSGFYYDTDNGESPVTAEDLPAIEAEMMKIVKENNPIVRKEISRAEALELFADDPYKVELITD
LPEDEIITVYDQGDFVDLCRGVHVPSTGRIQVPKLLSVAGAYWRGNSDNHMMQRIYGTAFFDKKDLKEFIKMREEAKERDH
RKLGKELDLFMVSQEVGSSGLPFWLPKGATIRRTIERYIVDKEISLGYQHVYTPIMADVELYKTSGHWDHYHEDMFPPMDMG
DGEMLVLRPMNCPHHMMVYKNDIHSYRELPIRIAELGMMHRYEKSGALSGLQRVREMTLNDGHTFVRPDQIKDEFKRTLEL
MVAVYADFNITDYRFRLSYRDPNNTDKYFDDDAMWEKAQAMLKAAMDELELDYFEAEGEAAFYGPKLDVQVKTALGMEETL
STIQLDFLLPERFDLTYVGEDGENTHRPVVIHRGIVSTMERFVAYLTEVYKGAFPTWLAPIQATIIPVSVEAHSEYAYEIK
ERLQAQGLRVEVDDRNEKMGYKIRASQTQKVPYQLVVGDKEMEDATVNVRRYGSKETSVEDLSIFIDSMAAEVHNYSR

Seq ID 239
VSNKKVAGTIILNLNDGSKKFLMHPIGEAIEFAMAKVSDEMTGLASMLQWFKEEVQLDVTSISLVELTNAHINKENVPLFV
FEMDEQALKNEMDKDYTWVAPAELKTIWSKYHIEGVPMF

Seq ID 240
MKKFDGVIFDMDGLLFDTELIYYTSTQKVADAMGLPYSKEVYLDYVGISDEEVQENYRRIYASYGHDTVEEFIRRSYDDTL
QEFRSGNVPLKPGVVEFLDFLDDQKIPRLVASSNVRPAIEMLLSHAGIQDRFVGIVSAEDVKRAKPDPEIFQKARQLLGTE
APKTLIFEDSFHGVSAAHSAGIPVIMVPDLLQPTEVIQEKTLHVLESLHQAPHYLK

Seq ID 241
MNELVFLHSQYIEEEPYTTDEVIAKYSQIKRESVSKLIKKYQKDLEEFGKVGPEIRAMESGQKAKFYQLNEEQATLLITYL
DNTEPVRRFKKALVRQFYDMKNGLYARRMERQKEKSVRKSMTDVIKELGLSPHYYKHYTDLVYKTALGFNAKQLREAREVS
KKSTILDFLTSEEIEAVNKREQQVATLLTLKMDYDTIKSILNGQGVLYQTTLKMPVATN

Seq ID 242
MPKPGETIHAIEHFTAGGGKGANQAVAAKRSGAETYFIGAVGNDGAGAMMTDLMSQDEINLTGVTTLEKTATGQAFIMVDN
AGENSIMIYAGANNAFTPKQVQEHQEIIEKSDPVIAQFESAIDSTIAAFKIAKKAGVKTILNPAPALEQVPEELLNVTDMI
VPNETETEILTGIKITDEASMRKAAEALHQLGIEAVIITVGSKGAFYDVNGRSGIVPAFKVKAVDTTAAGDTFIGALSSIL
EKDFSNLEEAIRYGNKASSLTVQRFGAQPSIPYQHELADK

Seq ID 243
MIKLAHIHKYYYSEEETLHVLDDINLQVDAGEFLAIMGPSGSGKSTLINLLGFIDKKFEGTYLFEDREIGDFSDKELSRIR
NEAVGFVFQNFSLIETLTVEENIELPLLYSGLTPKEAKDRVHEVLTKVGLPDKGKKHPKQLSGGQQQRVAIARAIVNRPSF
IIADEPTGALDSKTSEEILTLFQQLNNEGVTIILVTHDEETIEYCNRLIKVRDGKILEEVLT

Seq ID 244
MKKSGQKKRNKKVWFGIGAAVVVVGFIGAKTVFSSKEVEPEYTTYTITEMASLKLDGQVSFLDTRDIFFDPSLGKIAEINV
ENGKEVKKDSPLLTYNNSDIQATETEQANAVNRNNLQVQQAQENVNLATQKYNEALNKVAAAKQKLNTAKEAEEKETLNAE
IQQLNEAVSAANSEVAQANQALQLANSDAVGAANTLEQTRGKVNTVVTAPIDGQVTVDASAMSSTDKPVIKIATQKKNIQG
KVTEYDYDKLQTGEEVTVTTVGSGKSAPGKIVSIAQTPIAKNEGNPVVSYQFTVEGDFPWAEGLSTSIAVPQKQMIIPTAA
IQKEGQKEFVYVYKAGKAKKTPIETETNLGRKVVKSGLNWKDQVIANPNKELKDNQDVQVAAND

Seq ID 245
MIKKIIVVVAFMLTGFSLTAMSASAEEITDLFLQKEVTYSGVEGGKIGENWKYPQFVGEKAVDGDETTRWSADKQDEQWLI
VDLGEVKNIGELVLQLHAESPVYEILVSTDGESYQSIFKEENGKGGQPTKKYIDGNNVQARFVKYQQMKMWQHTNKQFYSS
SIISFEAYEKKRLPEAIKLLTENLTISEKRKQQLAFEVSPAGVDITEDQIEWSSSDPTIVTVDQTGNLTAVKSGEAKVTVK
IKGTEISDTIPVTVVAENKQYAEMRAKWKMRLLGTTQVDNDADVQQYRAQIATESLALWQTLNQAADREYLWERKPSDTVS
ADYTTQFTNIKKLALGYYEPSSELFEKPEVYDAIVKGIEFMIDTKKYNGTYYTGNWWDWQIGSAQPLTDTLILLHDDLLNT
DAEKLNKFTAPLMLYAKDPNIQWPIYRATGANLTDISITVLGTGLLLEDNQRLVQVQEAVPSVLKSVSSGDGLYPDGSLIQ
HGYFPYNGSYGNELLKGFGRIQTILQGSDWEMNDPNISNLFNVVDKGYLQLMVNGKMPSMVSGRSISRAPETNPFTTEFES
GKETIANLTLIAKFAPENLRNDIYTSIQTWLQQSGSYYHFFKKPRDFEALIDLKNVVNSASPAQATPMQSLNVYGSMDRVL
QKNNEYAVGISMYSQRVGNYEFGNTENKKGWHTADGMLYLYNQDFAQFDEGYWATIDPYRLPGTTVDTRELANGAYTGKRS
PQSWVGGSNNGQVASIGMFLDKSNEGMNLVAKKSWFLLDGQIINLGSSGITGTTDASIETILDNRMIHPQEVKLNQGSDKDN
SWISLSAANPLNNIGYVFPNSMNTLDVQIEERSGRYGDINEYFVNDKTYTNTFAKISKNYGKTVENGTYEYLTVVGKTNEE
IAALSKNKGYTVLENTANLQAIEAGNYVMMNTWNNDQEIAGLYAYDPMSVISEKIDNGVYRLTLANPLQNNASVSIEFDKG
ILEVVAADPEISVDQNIITLNSAGLNGSSRSIIVKTTPEVTKEALEKLIQEQKEHQEKDYTASSWKVYSEALKQAQTVADQ
TTATQAEVDQAETELRSAVKQLVKVPTKEVDKTNLLKIIKENEKHQEKDYTASSWKVYSEALKQAQTVADQTTATQAEVDQ
AEAKLRSAVKQLVKVPTKEVDKTNLLKIIKENEKHQEKDYTASSWKVYSEALKQAQTVADQTTATQAEVDQAETELRSAVK
QLVKVPTKEVDKTNLLKIIKENEKHQEKDYTASSWKVYSEALKQAQTVADQTTATQAEVDQAEAKLRSAVKRLTLKNSGEN
KKEQKNGGNNGHLNTSTGVDQTGTKQVKPSSQGGFRKASQFLPSTGEKKSIALVIIGLLVIASGCLLVFRKSKSKK

Seq ID 246
MKKSVLFTSLLVLSSLALAACGGGSDDKGASNGGSDNQVYTMVESQEMPSADPSLATDEVSFTTLNNVYEGIYRLDKDNKP
APAGAAEKATVSEDGLVYKVKLREESKWSDGKPVTAADYVYGWQRTVDPATASEYAYMFEPVKNAEKISKGELPKEELGIK
AINDHELEITLETATPYFDDLLAFPSFLPQRQDIVERFGKDYTKSSDKAVYNGPFTLTEFDGPGTDTKWSLTKNEEYWDKE
TVKLDKVAINVVKEAPTALNLYETGEVDDTYLSGELAQQMQNSPDLVQLKAASSFYLEMNQADEKSPLTNANLRRAMSYAI
DRDSLAKNILANGSLPSQGFVPVDVAKSPKTGEDFVKEAGSDKLVKYDKKKAVEYWNKAKQELGVSNLTVDLMVDDSEGAK
KMGEYLQGSLSDTLEGLKVTVTPVPMAVRLDRTLKGDFQIAVRGWSADYSDPINFLDLLESSTSNNRGRYSNPEYDKFIAA
SKTTDVNDPEKRWEDLINAEKTVIADMGVVPIYQKAESHLRAPNVKEIIYHPTGAKYDFKWAYKE

Seq ID 247
MSKKPSGYQKKVLTVLLVSFGFGGGLLAQFSPMKSGGFSSWFDDSYVKASABSSKTKEPAPVKIEKKVKPLSYGQQVNQEIE
KKQYDGHLDLPLELQTDAKWKDTAYGFGNVDKPNTIEINGCAIVSLAMVGSYMDHQEVTPLDVLAWAKNDFFMBGQGTAWS
IPSAYAEMKGYNCQEIGDIETVAAFLKEGHPVIISVKPGYFTTTGHIMVMSGVDEKGDFWINDPNDSEEKGHSKRTFTAEE
VMNEALNFWAFY

Seq ID 248
LKKSVLSALMVCSITLTSVALPSAAFADEYDTKIQQQDQKINALTSQMSDAEAKVAAIENDMVETAKQIDTLTAKKNKLSS
BVSKLYSEISDLNVRIQKREVQMTKQARDVQVNGQSDSIIDAVLDADSVADAIGRVQAVSTMMSANNELLEQQKEDKATVE
KKTKNVEKQIAELEAATKELNDKTESLKTLKIQQEVAKNDLEAQRSEEQGKKDGFIKQKKEABKRLAEEQARQRAAAKKAE
EQAAAQAQAAAQKAAAEQAKATKAANEAAASAAEEKAATPVVESSTTTESTTTQETTTSSTETESVVTTPVAAAPEKEKEV
PVTNPTTPEKGNEAKPGNGGVTSGKGQAAINAALADVGNSYATGWNQPGECLVSVRRWLAAGGINFGYGGPNSGYVASGATQ
VSWSNVQPGDVVQYESAYSPDSWIGGVHTVLVTGVSGGSVQIVEANNPGGSGYVSSNSNWSPAPPAGFRAVVWRFPG

Seq ID 249
MSSLLKRLVQLVLLVVAVLLIRHYVFSPAAVNGSSMEPTLHNNDRLWVTSIKKPQRFDIIAFPSPRNGQRVAKRLIGLPGE
TVEYRDDTLYINGVSLSEDYLASAKRNVSKNENYTQDFTLETLEATQSLTVPEGMYFVLGDNRPRSDDSRYFGFVKQASVE
GVLTFRYYPLDKIGFP

Seq ID 250
MANDDLQQRLDKGMYGTPLVNPEEQHKYMGTFRERCRLSMTVAEMKDAQNQKHLLEELAKHPEVTVLLNGEISSDLQSTYI
KLLNQHGANFKIVNNFVENNPDSLGLLLAEKHAVDEPVIDVTEKYPQATETPKEEPTAKKSFWQKLFHS

Seq ID 251
VYLKRIEITGFKSFADKTIIEFEDDVTAVVGPNGSGKSNITEAVRWVLGEQSAKNLRGGKMNDIIFAGSEGRKPLNIAEVT
VTLDNSDHYLALDYSEISVTRRLKRTGESDFFINKQACRLKDIQDLFMDSGLGKESFSIISQGKVEAIFNSKPEDRRGIFE
EAAGVLKYKQRKKKAEQKLFETEDNLSRVQDIIYELEDQLVPLAAQADAAKKYLALKEELTEIDVNLTVTEIQEAKAIWET
KTQELTAIEEKLAGASKQVHDLEGKLVRLRSKRNRLDEQIETEQQQLLQVTEALKQAEGQKNVLIERSKHTSQTASEYBET
LAETAEKIVRYREELQTLETAIAEKTAQRQTLKEALALATKDVEKYSKSSKELMEELRSQYVEVMQEQANTANDLKYLERQ
YQQETAKNQQSLAKHEALEBQMVEALAMKETLEKKEQKVAKQGLQEQLEEYTALKATLEAKRERLAQRQNDMYQAMNQVQQA
KARQKSLQEIQENYAGFYQGVKAVLRHKNQLTGIVGAVAELIEVPKEYTLAIETALGGAAQHIVVENEKDGRAGITFLKQQ
HSGRATFLPLTTIKPRSVSAMVQNRLAGAPGFVGIASELVRYPEQVQTVIQNLLGVTILAADLTSANQLAKLVNYQYRVVS
LEGDVMNPGGSMTGGANKRGNQGSLFSQAQELQTITEQMTQLETQLRSVEQEVQALSQEVKTATERAEMLRSAGEQNRLKQ
QEIDNKLANQTETITRLTKEKRLFEYESRELHQFLTEYQTKKATLTEQQANLTATKERLDAEMKQVEQEASQMETFKAQAQ
ERLTTVQAEQAVAAEQCAHFARQKQDKQEQLDELLIRETAIRQQLQQLSSHSSDHQLTEEGLAAQVAQLAEKQTALQTSLQ
TARSQRQALQEEVDELDTKLAEENKRQQQYLADKTQIEVLKNRAEMQLDSSLSYLQEEYSLTFEAAYEAYFPIDDLAQAQQ
TVKRLKQEIERLGPVNLSAIEQFEQVDERHQFLVSQRDDLLNAKEQLFETMDEMDQEVKERFKEVFEAIRGQFKVVFPNMF
GGGRAELVLTNPEDLLNTGIEIEAQPPGKKLQHLSLLSGGERALTAIALLFSIIRVRPVPFCILDEVEAALDEANVARFGH
YLSEFEDGTQFIVVTHRKGTMEAADVLYGVTMQESGVSKIVSVRLEEVKEGGAIEKSN

Seq ID 252
MQKVAKVIVDVPTMQTDQPFTYLVPENLNEQLAVGMRVEVPFGNGNRHVQGFVLAIEPVAATVLDETNVQLKELVAVLDLK
PVLNTEMLALADYMKEKTFAFKITCLQTMLPSVMRADYQKYIYLTDELSEELQDQLFYGLEEISWDQAQERGLLPQLMALR
KQQKVDIRYEVTTRNKVKMVRFIQAAKEFEQLEEIRLGLRKGAKKKEQLLYYLQRLGTEKVTAVKEMKELGFSTALLNEAA
KNGWLTPIEKEAYRDPFANQTFEKTTALSLNAEQQVAVETILQSVQEQQSQTYLLEGITGSGKTEVYLQVIAEVLNQGKTA
IMLVPEISLTPQMVQRFKSRFGEHVAVMHSGLSQGEKYDEWRKIERGEAEVVVGARSAIFAPIENIGVIIVDEEHEASYKQ
EETPRYHARDLAIWRSEYHHCPVVLGSATPSLESRARAQKNVYQRLRLTQRANQAATLPTIDVVDMRQEVENGNVSSFSMS
LQEKLQERLEKNEQSVLLLNRRGYSSFVMCRDCGYVLPCPNCDISLTLHMDSKTMKCHYCGHEERIPYRCPNCGQDKIRYY
GTGTQKVEEELQTLLPDSRILRMDVDTTRRKGAHEKILRTFGEGQADILLGTQMIAKGLDFPNVTLVGVLNADTALNLPDF
RSSERTFQLLTQVSGRAGRAEKPGEVIIQSFNPEHYAIQLAKAQDYEDFYTKEMYIRHRGDYPPYYFTVQITASHPEENEA
AKQMFQIATKLKQGLSPQAILLGPTPNAIMRVNNRYFYQVIIKYKQEPMLQPLLKEILTDTQRATARGLKLSIDAEPMNFI

Seq ID 253
MDLHLTVDLKKTKTLIDLADQAHLIIEQTTKLPKKEAAILFLLIDNQLVALGLAEEKATYKEVSFDHSILLKPTWDTELAY
LAQAFLDDAKESGLTLEATPTTVKIPKSAVATVTNYKETVTFILERFGYSLFKKPAPKKARPAKARHKWTKEVSQIPFYID
TRQSKATVFWQKRNEMLIKAGATMMPEAPLNKDGSVGFSARFGEKLRDERKGQFKDFVTTEDIVLKSVNEVGLFLYFAGTN
SWLELVDENGKTLNEWTVVE

Seq ID 254
MNRWKVYATVIACMLFGWIGVEAHASEFNPAVTPTIPENQVDKSKTYFDLKMAPGAKQTVEIQLRNDTDEDITIENTVNSA
TTNLNGVVEYGQNGIKPDKTLRFNLKDYVEAPKEIILPKHSQKTLPLTITMPKDSFDGVMAGGITLKEKKKKETTTSADQSK
GLAINNEYSYVVAIILQQNETKVQPDLKLLGVKPGQVNARNVINVSLQNPQAAYLNQLHLINTVSKGGETLYQSDTEDMQV
APNSNFSYPISLKGERLTPGKYVLKSTAYGVKDEKGTYQVKGANGEERYLYKWEFTKEFTISGDVAKELNEKDVTIKGTNW
WLYLLIALIILALLLLIFFLYRKKKKEEEQQSEQ

Seq ID 255
MSKNFWATLPKPFFVLAPMEDVTDVVFRHVVKEAGAPDVFFTEFTNSDSYCHPEGKDSVRGRLVFTEDEQPMVAHIWGDKP
EFFREMSIGVAEMGFQGLDINMGCPVPNVAERGKGSGLILRPEVAAELIDAAKAGGLPVSVKTRIGFTEMAEMEAWITHLL
EQDIANLSIHLRTRKEMSKVDAHWEVIPQIMAIRDRVAPQTTITINGDIPDRQKGLELAEQYGVDGIMIGRGIFKNPYAFE

KEPKTHTPQELLGLLRLQLDLQDKYAELVPRSIVGLHRFFKIYVEGFPGASDLRAQLMNTKSTDEVRQLLATFETEHGVLD

Seq ID 256
MRYLDAEAIENIATGAAFLGTGGGGDPYIGKMMALSAIEENGPVELVSPEEIAAEDFFLPAAMMGAPSVAIEKFPKGDEFV
RVFEKLGKYLDQETIAGTFPMEAGCVNSMIPIVVAAKLGIPLVDCDGMGRAFPELPMVTFHLNGMSATPMAITDEKGNIGI
METIDNTWTERLARVQTVEMGASALVSIYPATGKQLQDYGIHNIVTLSEEIGKVIRGTYADEQEKRQALVEVTDGFELFQG
KILDVEREVKGGFNLGRVKLSGLNSDAGSEAVVHFQNENLIAEKDGQVIAMTPDLICMVDLETLTPVTTESLKYGKRVQVM
GLKANAAWRTKKGIETVGPRYPGYEMDYQPLENLVAKEDK

Seq ID 257
VTIVLLLLLNVSIMFLLGQLLSENSPLAGVIWATKDAVAKELNTDLFWSWNKFVLPIFFLVDAAVLYWRLIRRYHQMQLRH
IISELHYIADGNYNHRIPFELSGDLAKVVTSINGLVDSTVAAIEDERRIEKSKDELITNVSHDIRTPLTSIIGYLGLIEDR
QFHSQEDLLKYTHTAYVKAKQMKLLVDDLFEYTKVRQPSVPIHTTTFDMAQLIEQLAADFELEAKKINMQIQVKANPASLM
MEGDTEKLVRVFNNLLSNALKYGKGGHHIVMEVDKVGTEAIIAVRNDGPAIPKHSLDQLFDRFYRVEESRSQETGGTGLGL
AIAQSIVALHGGYIYAKSDQKWTSFIIHLPLQRTNKKSES

Seq ID 258
MRAEALRAIVAENRQLEQNLTKRNEQYIFDLKKSLKAANLSEEELALALHGILPELVAGQKTGKTARQLFGTVSERTEAIL
NKPAEVKEPAGWMIWLDNTLLLLGLLTIMLAAMSLFSKGTAQPLGLTTYILGAMAGGYVFYLMHKYVYRFDRQGGDKSKRP
GWLKTTAILFGGMFLWIAVPAGSAMLPPVINPILDPMIALVIGALAFVARYFFKKKYNIQGSFMTRQ

Seq ID 259
MNLVLAEIAAMGLKNISIAPSSIANVHEPLIEHIKNGVVTNITSSGLRDKVGAAISAGIMENPVVIRSHGGRARAVAAGDI
HIDVAFLGAPSSDAYGNANGTKGKATCGSLGYAMVDAKYADQVVIITDTLVPYPNTPISIPQTDVDYVVEIDAIGDPDGIA
KGATRFTKNPKELLIAEYAAKIITHSPYYKEGFSFQTGTGGSSLAVTRFMREQMLKDGIKASFALGGITNAMVELLEEGLV
EKIIDVQDFDHPSAISLGENANHYEIDANMYASPLSKGAVINQLDTAILSALEVDTDFNVNVITGSDGVIRGASGGHSDTS
MACKMSLVIAPLVRGRIPTIVEQVNTVVTPGTSVDVVVTEVGIAINPKRTDLIDCFKTLDVPQFTLEELKDKAYNIVGTPE
PIKYGDKVVALIEYRDGSLIDVVRNV

Seq ID 260
MTMNPFKGKQFQQDVIIVAVGYYLRYNLSYREVQEILYDRGINVSHTT
IYRWVQEYGKLLYQIWKKKNKQSFYSWKMDETYIKIKGKWHYLYRAIDADGLTLDIWLRKKRDTQAAYAFLKRLVKQFDEP
KVVVVTDKAPSITSAFKKLKEYGFYQGTEHRTIKYLNNLIEQDHRPVKRRNKFYRSLRTASTTIKGMEAIRGLYKKTRKEGT
LFGFSVCTEIKVLLGIPA

Seq ID 261
MNKKRIIFGLLSFFLPIFLVFGGLLFFLLLLTSTSDTSKNDCIQPSINNPTDATDTPKSIEQFVKSHKDAYLLSWKAGGFL
PSASISQTMVENGFNPTNPSGTSFWQAHNMGGVKTSKKEDFPVTLATFGQDSVDISGTKPGSNVGDGTGGAYTWFKDYNAG
IVGKAEPMAHQTLYTGAINNTDGLSTLSAIYSGGWATDPTYLMKLQATYNSLGKQFQWLDQEAIQKYGNAPFKKSELVLNI
PGKSPITNEKYGKNSDCVVTSDTSDQVTGQNTAPSLEVPSAYKGKLTLPPIDSNDYAGNNYPFGQCTWYAYNRMAQIGKPI
EWFSGDGGNGAGWANSARAKGYTVVKVKPSVGWAASMQGGIGGSAPPYGHVAVVEYVNSDGSILVSEANVINQGSGTRSWR
VLDRATVEQIDFIQGKGA

Seq ID 262
LNLQIKIKMEEIMMKKNLFASLFSATLLFGGSEISAFAQEIIPDDTTTPPIEVPTEPSTPEKPTDPTPPIEPPVDPVEPPI
TPTEPTEPTEPTTPTEPEKPVVPTEPTTPTEPTTPTEPSEPEQPTEPSKPVEPEKPVTPSKPAEPEKPVTPTKPTEPEKPV
QPAEPSKPIDVVVTPTGELNHAGNGTQQPTVPIETSNLAEITHVPSVTTPITTTDGENIVAVEKGVPLTQTAEGLKPIQSS
YKVLPSGNVEVKGKDGKMKVLPYTGEEMNIFLSAVGGILSVVSGFVIPKKRKAKV

Seq ID 263
MKQQTEVKKRFKMYKAKKHWVVAPILFIGVLGVVGLATDDVQAAELDTQPGTTTVQPDNPDPQVGSTTPKTAVTEEATVQK
DTTSQPTKVEEVASEKNGAEQSSATPNDTTNAQQPTVGAEKSAQEQPVVSPETTNEPLGQPTEVAPAENEANKSTSIPKEF
ETPDVDKAVDEAKKDPNITVVEKPAEDLGNVSSKDLAAKEKEVDQLQKEQAKKIAQQAAELKAKNEKIAKENAEIAAKNKA
EKERYEKEVAEYNKHKNENGYVAKPVNKTLIFDREATKNSKVVSVKAAEYIDAKKLTDKHKDKKLLISMLSVDSSGLTTKD
SKKAHFYYNNGAGGTLVVLHKNQPVTITYGNLNASYLGKKIASAEFQYTVKATPDSKGRLNAFLHDDPVATIVYGINIDPR
TKKAGAEIEMLVRFFGEDGKEILPTKENPFVFSGASLNSRGENITYEFVKVGNTDTVHEINGSKVARHGNKVYSKTDIDVG
TNGISISDWEAVQGKEYIGATVISTPNRIKFTFGNEIVNNPGYDGNSMWFAFNTDLKAKSITPYQEKGRPKQPEKATIEFN
RYKANVVPVLVPNKEVTDGQKNINDLNVKRGDSLQYIVTGDTTELAKVDPKTVTKQGIRDTFDAEKVTIDLSKVKVYQADA
SLNEKDLKAVAAAINSGKAKDVTASYDLHLDQNTVTAMMKTNADDSVVLAMGYKYLLVLPFVVKNVEGDFENTAVQLTNDG
ETVTNTVINHVPGSNPSKDVKADKNGTVGSVSLHDKDIPLQTKIYYEVKSSERPANYGGITEEWGMNDVLDTTHDRFTGKW
HAITNYDLKVGDKTLKAGTDISAYILLENKDNKDLTFTMNQALLAALNEGSNKVGKQANSVYLEVERIKTGDVENTQTENY
NKELVRSNTVVTHTPDDPKPTKAVHNKKGEDINHGKVARGDVLSYEMTWDLKGYDKDPAFDTVDLATGVSFFDDYDETKVT
PIKDLLRVKDSKGEDITNQFTISWDDAKGTVTISAKDPQAFILAHGGQELRVTLPTKVKANVSGDVYNLAEQNTFGQRIKT
NTVVVNHIPKVNPKKDVVIKVGDKQSQNGATIKLGEKFFYEFTSSDIPAEYAGIVEEWSISDKLDVKHDKFSGQWSVFANST
FVLADGTKVNKGDDISKLFTMTFEQGVVKITASQAFLDAMNLKENKNVAHSWKAFIGVERIAAGDVYNTIEESFNNEKIKT
NTVVTHTPEKPQTPPEKTVIVPPTPKTPQAPVEPLVVEKASVVPELPQTGEKQNVLLTVAGSLAAMLGLAGLGFKRRKETK

Seq ID 264
MKYERPLKRESQIKEFELGTHAAVIEKVQKKRSQKGNDMFLLSLLGKSNEKGVYFLTFGNDYTEDNLRYILASIQDNGVEI
PDVDFGYNRETFEFLKGKDVYIQVEEQEYKGKVKHAVTNFLTQDEFEESEEMEFSESNTEEDW

Seq ID 265
LVEQREIVHRYYVLGTSTLDKVLKVLLNLSSDGVLALKDKFLPLKGENNLYKLMNREDPLTSAQLHEKVNLNKLKEQLEAQ
GLPFAFKETKEGTNFYFRVKDTELAKKALERVLTDIKKNPQMILRKPNTMTFDERLAYTAANKKYVGKIDQTKAITKGRKL

Seq ID 266
VRMNKKTPKPLTPPKAFFNWCTAQIPTYEWQNKKETILASSRKNCPIIKKRLTKYSRLSFPTKFYSFGIILVRAKRIEIQT
HSYWQTITDGIENLIYEPSNLERFSNDTHVKAHYENGHWHEGLLANYGFMSSAYTNTVFYPNNWQEKLTTVSELKYLQLPE
IERQELAHIYKRNEIEFLQKIGATTLANEIIFDDYRNVFGLYFHKVDMRVITKKWLKANKQKLKTRNPTFHEFMLEKTLK
ERNAPMINEIEKYVHYSQIKQLPKEVNLTKFQKWFIRKGERFDYYMDYLHMLEELNTFLNNDSVLYPENLQVAHDNAMNTL
NLLKSEIEEKQYQERKNQIKALEAEIDDLLFLTPHSLQEIIQEGSILRHCVGSQHYIERHTQGKTTIVFIRRKEKPDMPYF
TLEYRNQQVIQIQGKCNRQEVPEKIKQAVRQWQENLQHALSS

Seq ID 267
PSEKGNQGRVEKRSDAKSTNHVNAGRYTE

Seq ID 268
ENEFKLHKIGLHSFLDRVHQSVVLGVRSLRSIKKKCNHYPQTEKQKQDKAGQAMRMRQGIVLGMCLIVLHS

Seq ID 269
NIFSKSSAPIIRRFIRSMARWDVCFLDDRSLDAGASFDGFCLPNVCYRYVDWDVNYGSWWYGNV

Seq ID 270
RGRCYIFNEVPRCWQENSPTNDFRFKRKIRRIRVFRSSGCGHDRNRSGDNKQSSIGRSVRSIKCPRLQ

Seq ID 271
TRLWGRQLCRKSQIRKSSGINYRR

Seq ID 272
LGLSFIKHVLAYGYVLWGNIHKRQIHWELMST

Seq ID 273
HLVFVHVKDVNQPGKSFLTRNGLLKLSNLLIGIQRNH

Seq ID 274
KTVVARQLRWKWRLQKQGLHRNKLGISMLMVRQPQRMTLRKQRPSNVSLVSEHRKFQFQVQRV

Seq ID 275
CYSKFRQTSKQWPNKRRQGD

Seq ID 276
VASPSKSAPFFTHAKKCQNQRRNSCFSFSSPSWRLCHLNV

Seq ID 277
NKGRQPIYRKLQRAPLILKYLSPAQVNQLIHNQLQNQLATKWQVVAKKFLGKQVIRSNIKHGALVQSFV

Seq ID 278
ESIKDFNNVSRKKQIKRVQSKWKIIFYSVVISLDSAVKPIGKFSDIFKEALFSNTTPKTKKSSVKEYVKI

Seq ID 279
HLWLVLRYLPKQNQASKHLLMKSEIQQQPLPIAMICMLPS

Seq ID 280
VMFRQNHRRDKKKKRRSEQRRKKPKLQNAKQKLKPLPNKNYKSGNAWNKRQPSA

Seq ID 281
RIANWNLEYTNNLEVILLGKKGSKIKKKKRRLKEKAIANGTYSKRGKSDGMYKMQGTDDCLGKR

Seq ID 282
KTNTALAGMATLAMMNGTTAKASSCPAPTIFPVAIMIRLR

Seq ID 283
RLPAPPLRKTRSLPSMPKSWNRCSPKT

Seq ID 284
GGGTHLKQYEIILADPPWRYSQKGGAGSGRTSLRHHEHGGFMLTAYQPFCCAGLRAFSLEYIPTASRSTASDSCMGIHL

Seq ID 285
RKQLITVVYKSKIQRFMSAIHGNQKRTLFQQQIKQVKTFPSKKSLFQVKLITPKQAFIQLFTVTKVKKKQPM

Seq ID 286
QELRIVVIQIQINQKRIRQQQMQAKKLKKYLILMLKLAIPSK

Seq ID 287
LSWQSSASRRCFGNRDKSIPSSRWRKLSNVPT

Seq ID 288
RNTQGKCRFSLRVSLFFGKVCRKVGSFAHELSIDFCRKPTNRNRTLDFKARHISRCRGYVRICFR

Seq ID 289
FPSGLLFARGNVGAGSASGDSSIRGTVEHATFNGSRTSGGTESARSI

Seq ID 290
NQRKKFNQNQNQSPLRKSQNRNNLRQPLRKIKKLRFLLKKKSLHPRKKTNLWRNRQCQKKKRNKQKKQRLIPYK

Seq ID 291
QQVCLAPLSHSLRAWESSNANVKTSSLL

Seq ID 292
AVMHDQKTNRQPLTQSTKKKPKNKQNMRSMVSLKILQNGTSGKLTMENLKIYYKKRNLLLTTKDF

Seq ID 293
PNSCVYHMNDQRIKFQATKNKIGPKMLPPIAKEAPKAPNIGRTVNQKTGSPTPAKGPIIPTLTP

Seq ID 294
IKCPTSRQLKHFFFHAFRQTCIPKAFNHINNGDYFFIISHFASVQIIRKVF

Seq ID 295
RHISNGGTFLAVTEMAIDDFDSRTFQIF

Seq ID 296
STLSIIALIFSKHCVVCSLIPPATNAFVLGSNGS

Seq ID 297
NKSSAFSPAKRCPPLTTTCLAPICTIRFAASSISATEEISPKPDKMPASIKLGVATNANGSKSFFRPSTLLGCRICAVPLQ
AMTGSITTF

Seq ID 298
LPLVVYNESNRKMDALENSCRQSRLSRQKQRPEWLFQLPLRIVLLHTK

Seq ID 299
TSNDSIMIAKLSNSFFKKFRSRTKQLIAASDCRTKTNNVASNPRIVDQLAPNVPQRMVKKPVVSNFTNKPTPAATKGPLKR
LRLCSILW

Seq ID 300
SALMSTTRPSGLSNKVTIFTDAALRYLNKSTKYANVRPLSMISSTIKICLPSTFVSKSLTIFTSPEETVALP

Seq ID 301
HVLQKYQRPQYLRYPRKIEPFPQFVICFLEFQPIQNL

Seq ID 302
MSIITPADTFLNRLINVAADSFKPFLSNVPVANIIIIKIAMPRLFTKVVPSVSSIDVKPSPPVKAVTIKATITIKIASNFN
AKPTMTINTPNKGNKSIKSLLLKIKFHFPLYANFRSFVRLF

Seq ID 303
RLNRSVLSLVPLLAVLGDKQKNHLAFFVKAAAVSSLIVHH

Seq ID 304
SVCIARHLPHLDELGHHKTNHHGRYYN

Seq ID 305
VAPNKVSGRVVNTVKLSSVPATLKVTLAPSERPIQFRCISLILSGQSTSSKSSNKRSA

Seq ID 306
MIPASIKMTRSATSFAKPISWVTTTIVIPSIAKFFITDSTSPTISGSKAEVGSSNSITSGRMAKARAIATRCCCPPES

Seq ID 307
HNPMKIEAKNAGIASSKSVQLMSLKEANIIIPTKINAGAVACAGTIAI
SGEKRLLRANRIATTTDVKPVRPPAPIPAALST

Seq ID 308
YPFVISTKYSSNNKFSSAGFFSYASVIFCKNCARIIQPARQIFAISPNGKSHPYSSEAARSCAKPCA

Seq ID 309
PSSSINVTFVVVEPASIPKKVTPVC

Seq ID 310
AIINDVRPFRSCCSPSCNNSSVFVSILEVASSKIKILGSASKARANEINWRCPAESVLPRSSTTVS

Seq ID 311
SSSSGNTCGQKDSASQNNCSVTVPLNLGSSSSAKGKLATSAAV

Seq ID 312
IIPASTGECKASMPNIVKANAIANVYNKLMMGLNNNEFIRIIATRVFK

Seq ID 313
TNCIATPQAIPNFKSPKIMPQTNPPTKGRPIV

Seq ID 314
YPFQIDKSTPWSPFAANVLIGDTKNKTTSNVANSFFF

Seq ID 315
PFNDLRISIITFSLDISMPTKGSSKKITSAPWAKTRAINTRCNWPPESSPIWRCA

Seq ID 316
FLHKSHCPVEPKKTVPVPSLIC

Seq ID 317
DTNCCANSVAVTPTSSNRTMIV

Seq ID 318
GSPSSSYRVSTITSRVRVYCREKMRSISGWTSAKVKALNAVER

Seq ID 319
LLCHYCFFLIPQCDRNNRDSLCHVLGS

Seq ID 320
KTIAKVNNATNQFRAEPKSALPAPPAIYLIATGNKDSPMLKITVPVTKGGNNTRIFLMNKPTIITRILPTNWEPKIAAIPK
SAPMAIKIGTYAKLAPIITGNREPTFPKIGKS

Seq ID 321
PTLILKMKPFGSPSSNSLLSNFKTNHALTYSS

Seq ID 322
GTLSITVAPVCTMTAAESEDVPISKIAGPKEPVAKITAVDVEPTAKPIKIPEIGAKKVPEIPITEIKWTICVDKSETEINP
AIAPAPIKIKTVSVTLFKPEDA

Seq ID 323
TKSRANVDTLQEIYTIFLGGCFKRAVNNLRSKPVRGGSTTMASGLKPSLNHLGNKSSAVPTA

Seq ID 324
LSCETKKSTGPSLLTVCILIVYTKTITPIPLKAVAINKVKPVSVLLLLNK

Seq ID 325
FLFSLPVRREPPLKFHKQSSVLAPFVPITTIILQNQHLQQKLVLNRQRHTFG

Seq ID 326
NKPAKNAPKTAPVNANGGIPKNVIRIKPQTSAMSVP

Seq ID 327
INSQIAPMPINCLNLCADIIKSNHIQNNMHKSNVKKHWRDKPPVFMILSYQWTNHCSPIN

Seq ID 328
FSFSLYEFPHKKRQIIQQDRRR

Seq ID 329
AKNALIPKAGAVAIGYFANTPIKIQPKALAKAVAVKTAPSSMPEAPSTVGLTNKIYAIEINVVKPANISRFIFVFNCSSSK
YELIFSFMFRPPKI

Seq ID 330

LFSDTFSISLTAFVSLNSIIGNSRRNSDCKINEELIPLGQTISAFSFSCRCSFFCPKISPPNNMVKKPISSDGIDRFVNFS
KGANACSQR

Seq ID 331
INSHSIQRVIRVPCPKSSSGSKKSSSVS

Seq ID 332
VNQCLKRQIYRHIFLIHHSSGSCALLPFCNNFPYKINAPNKLNHHTERGIQMKDLFNHQLTKHQ

Seq ID 333
SAISIFSVPELMYKVTDRPFFTDDFASGSVLTTLPSSILSSGIFLVSLSTFSPALFSACEASAFVSFVTSGTSTSLPPDIA
KAKPPIAKEINAAIAKPIHNFLRFFFGGCSAKMSSAISSGTSTGVVSSGSAMSSGIGVNTFVSPNKAVGSHAGSFKRLADK
DVYKSSAISFAVL

Seq ID 334
GSPPPPVPKKAAPVAIFSIASASK

Seq ID 335
SPAITAPNAKEPTVIERPDIKLPAV

Seq ID 336
IRSRNDRLASGTLISFSVKCNTSANKMEKPMPNNIQPLMNGIFASGMLVLAPIAVTITPKKAPGFEIS

Seq ID 337
MFVNIGMNPIFIDRKAIKTDTIGATAIGINNKGLKITGSPKIKGSLMLKNAGTNEIFDMAFILDECPTMAIAIAKPIVQPA
PPIVAN

Seq ID 338
QANSVRPMNSHAHLQADSNNPLLHVLD

Seq ID 339
VKLPVPSFEISCAFCSRLAHPDNSKVAAVSSSHFFPINSF

Seq ID 340
VIQQESRLRVADNSGAREILTIKVLGGSGRKTANIGDVIVATVKQATPGGVVKKGEVVKAVIVRTKSGARRADGSYIKFDE
NAAVIIRDDKSPRGTRIFGPVARELRENNFMKIVSLAPEVL

Seq ID 341

```
atggtaggga ttatcctagc aagtcatgga caatttgctg aaggaatctt acaatccggt      60
tcaatgatct ttggcgaaca agaaaacgtc aaagccgtga ttttgaaacc aagtgaaggt     120
ccagatgact tgagagccaa attggaagaa gcggttgctt cttttgataa tcaagacgaa     180
gtgttattct tagttgattt atggggcgga acaccgttca atcaatcaaa cacattattc     240
gaagagcaca aagataaatg ggcaatcgtc agcggtttga atttaccaac gttgattgaa     300
gcctacgcat cacgctttte aatggaatct gctcatgaaa ttgcagcaca catcattgaa     360
acagcaaaag acggtgtaaa agtaaaaccg gaagaattag aaccggcaga agcaccgaaa     420
gcagcagttg aagacgcaca acctaaagga gcacttccag aaggaactgt ggtcggagac     480
ggaaaaatca aatatgtatt agcacgtgtt gactcacgtc tgttacatgg acaagttgcg     540
acagcttgga caaaagcagt gcaaccaaac cgtattatcg ttgtttcaga tgcagtatca     600
aaggatgatc tgcgtaagag actaatcgaa caagctgctc ctccaggtgt aaaagcaaat     660
gttattccta ttagcaaaat gatcgaagta gcaaaagacc ctcgtttttg aaataccaaa     720
gcattgcttt tatttgaaaa tcctgaagac gtattgacag ctgttgaagg tggcgtagac     780
atcaaagaat taaatgttgg gtcaatggcg cattcggtcg gtaaagtcgt tgtaagtaaa     840
gtattgtcta tgggacaaga agacgttgaa gcgtttgaga agcttgaaca aaaaggtgtg     900
aaattcgatg tacgtaaagt acctaatgat tcacgcgata acatggatga tattttgaaa     960
aaagcaaaag ctgaattagc aaaagcgtaa                                       990
```

Seq ID 342

```
atgattatta aagaagagat ttttgtacct aaagaattag tgaaagaaaa aattgacctg      60
ctcgttgaaa acctgacaac gatcaaagac gataatggtg aatttttgtt ggactttgat     120
ggactgaaag tagatgacaa aagctggaca gtgtggaatt ggccacaagg tgtcggtcta     180
tatggtattt ataaaaatta ccagatgaca aaaaatcctc gcgcttatca agtggtcaat     240
gaatggtttg aggatcgtat ggaagaaggc gctccgccca aaaacgtcaa tacaatggca     300
ccgctattga ctatggctta cttgtatgaa gacacgaaag attcacgcta tttgccttat     360
ttagaacaat gggcagaatg ggtcatgaat gacatgcctc ggacaaatga agacgggtta     420
cagcatgcca catatggacc ggaaaataag aatcagttgt gggacgatac attgatgatg     480
acagtgcttc ctttggcaaa aatcgggaaa ctgttgaatc gcccagaata tttagaagaa     540
gctcgttacc agtttatgat ccatatcaaa tacttgatgg ataaacgaac aggattgtgg     600
tatcacggct ggacgtttga aggaaatcat aactatgccg aagcattttg ggctcgtgga     660
aactgctggt tgacgattgc tattccagag atcattgaaa tcttagaatt atcgaaagaa     720
gatgcattgc gcaaattatt gatcgaaaca cttgaagcac aagtgcgtgc attaaaaact     780
tatcagtctg aaagcggttt atggcatacg ttgttggatg atccttcttc ttacgtagaa     840
```

```
tcttctgcaa ctgcaggatt tgcttacggt attttgaaag cggtccataa acgttattta    900
ccacaagaat ataaagaagt cgcttatcgt gctatccaag gattgcttga acaaatcgat    960
gaaaaaggag aagttcaaaa cgtttcaatc gggacaggaa tgggggacag tttagatttt    1020
tatcggaata tcggaatcac agcaatgcct tatggacagt cgttaacggt gctttgtttg    1080
acagaattgc tcgtctctta ttgttaa .                                     1107
```

Seq ID 343
```
atgaaaatag caattgcagg cgctggcgca atgggaagtc gctttgggtt aatgcttcac     60
caaagtggga acgaagtatt attaatcgat ggctgggcag aacatgtcca acaaatcaaa    120
gagcatggat tacaagcaaa ttttaatggt aaagaggtag aagcaaaact accaatcgtt    180
cttcaatccg aagtagaaaa agaagatcaa gttgatctga ttattctatt taccaaagcg    240
atgcagctgg aaaaaatgct gcaggatatc caatcattaa tcaaaaaaga tacagaggtc    300
ttatgtctat taaatggtat cgggcatgaa gatattattg agaaatttgt accaatggaa    360
aatatctata ttggaaatac catgtggacg gcaggtcttg aaggtcctgg tcaggtcaaa    420
ttatttggaa gtggttcagt agaattacaa aatttaggtg atggaaaaga agcagctgca    480
aaaaaattag cagacaaatt gtctgaatca ggattaaatg ctcatttttc tgacaatatt    540
cactattcta tttatcgcaa agcatgtgtg aacggaacaa tgaacggact atgcacaatc    600
ttagacgtta atatggctgt gctaggaaaa acatcgactg ctcataaaat ggtggcgacg    660
attgtcaatg agtttgccaa agtagcagca gtagagaaga ttgaactaga tgtcccagaa    720
gtcattgcac attgtgaatc ttgttttgac ccggaaacaa ttgggttaca ttatccttca    780
atgtatcaag acttgattaa gaaccatcga ttaacagaaa tcgattatat caatggcgca    840
atttctagaa aagggaagaa atatggtgtt gcgacgcctt attgtgattt cttgactgag    900
cttgtccatg caaaagaaga tagtttgaac gtaaaataa .                        939
```

Seq ID 344
```
atgacttttc cttattcaga aggaaataaa agatatcatt cttggaatta cgctttacga     60
aatgagtttg gcgggaaaat tttcaaggtg ccgatcgatg gaggatttga ctgtccgaat    120
cgagatggga ctgttgccca tggcggttgt acttttttgca gtgtatctgg ttcgggagac    180
atgatcgttg cacctgaaga ccctcttcca atccaattcc agaaagaaat cgacatgatg    240
catcaaaaat ggcctcatgt ccaacaatcc atcgtttatt tccaaaattt tacaaacacc    300
catgcgccag ttgaaataat caaagaacga tttgagcaag tagtcaattt accaggtgtc    360
gtaggattat cagtcggtac acgaccggat tgtctgccag atgaagtagt tgattacctt    420
gctgaactga cgagcggat gtatttatgg gtagaacttg gattgcaaac aacttttgaa    480
gaaacttcaa aactgatcaa ccgagcacat gattatcaaa cttatttaga tggcgtggct    540
aaaattacgca aacataatat ccgtgtttgt acgcatttga tcaacggatt accaggcgaa    600
agtttagaaa tgatgaaaga aaatgtcaga cgcaccattt tagattcaga tatccaagga    660
atcaagctcc atttgatgca tctgatgcgc aaaacgagaa tgttgcggga ctatcatgaa    720
ggtcgccttc aattaatgag ccgtcctgac tatgtcaatg tcatttgcga ccagctggaa    780
atgataccaa aagagatcat catccatcgt ttgacagggg atgcgccgtg ggattccttg    840
attggaccaa tgtggagctt gaaaaaatgg gaagtcttga atgcaattga tgaagaactt    900
cttcgcagag atagcttcca aggaaaatat gatgtacgga aaaaggtgag cgtctga      957
```

Seq ID 345
```
atggcaaaaa caggaatgta tgtaggcctt gatattggaa cgacatctgt taaagttgtc     60
gtggctgaat atatcgacag ccaaatgaat attattggag taggaaacgc aaaatcagaa    120
gggattaacc gaggcattat cgttgatatc gataaaacgg tccaagcgat acaacgagca    180
gtacgacaag cagaagaaaa agcaggtatc cagatcaaag gcgtaagtgt cggttttgcca    240
gcaaacatgc ttgaagtaga aaaattgccag ggaatgattg cagttaacgg agattcaaag    300
gaaatcaccg atgaagatgt cagaaatgtg gcttcagctg ccttgttcg ttcaatacct    360
cctgaacgtc aaatcgtgtc aatattgcca caagattta ctgtcgatgg gttcgaggga    420
atcaaagatc ctcgtggaat gatcggcgta agattagaaa tgtacggtct gttgtttaca    480
ggacctaaaa cgatcatcca taacatacgt aaatgcgtgg aaaatgccgg attaatcgtc    540
aacgaaatgg tgattactcc tttagcgctg actgaatcaa ttctttctga cggagaaaaa    600
gattttggta cgatcgtcat tgatatgggt ggcggccaaa caacaacggc tgttatgcat    660
gataagcaat tgaagtttac caatcttgac caagaaggtg gcgaagttgt cactaaagat    720
atctctattg tattaaatac gtcatttaac aatgcagaag cgttgaaaat aaaactacggg    780
gatgcatacc agaacggac atcaccagat gaagaatttc cagtagatgt gattggtcag    840
tcagaaccag tgaaagtcga tgaacgctac ttgtcagaaa ttatttcagc tcgaatggaa    900
cagattttca acaaagcaaa agatgcattg gaccaaatcg aagcattaga gttgccaggt    960
ggcgtcgtat tgacaggcgg agctgctagt cttcctggtt ttgtggactt ggcacaagaa   1020
atcttcggcg taaatgtaaa attgtatgtt ccaaatcaca tggggctgag aaacccggta   1080
tttacaaatg taataagtat cgtggattat tcagcaaact tgagtgaagt gtatcaattg   1140
gcaaaaattg ctgttacggg tgaaacctca gcagctcgtc aaatggtagt tgaacaagaa   1200
acaacaaaca catatgagag ctatgaagca cctgaagaaa caatctacga tgagccggaa   1260
ccaaaagagt ccggagaaaa tgtaaaaaat aaaatcaagg gtttctttac aaatattttt   1320
gattaa                                                              1326
```

Seq ID 346
```
atggcaaaaa aaacaatcat gttagtttgt tcagcaggga tgagtacgag tttattggtc     60
acaaaaatgc aaaaagcagc agaagaaaaa gggatggaag ccgatatttt tgcggtctct    120
gcctcagatg cagataataa tttagaatct aagaatgtcg atgtattatt attaggtcct    180
```

```
caagtacgtt tcatgaaagc tcaatttgaa caaaaacttg cgcctaaagg tattccatta    240
gatgtcatca atatgcaaga ctatggcatg atgaacggag aaaaagttct tgcacaagca    300
gaaaatctaa tgaaataa                                                  318
```

Seq ID 347

```
atgaaaaatc ataaaaaaat aaacgttatg ttaggagtcc ttttccttat tttaccatta    60
ctcacaaaca gcttcggcgc aaaaaaagtg tttgcagagg agacagcagc tcaagtcatc    120
cttcataaaa agaaaatgac tgatttaccc gatcctttaa tccaaaacag cgggaaagaa    180
atgagcgaat tcgatcaata ccaaggatta gccgatattt cattttcagt ttataacgtc    240
actcaagaat tttatgcgca acgagataaa ggagcgtccg tggatgcagc aaaacaagca    300
gtccagtctt tgactcctgg tacaccagtt gcttcaggaa cgacagatgc tgatggaaat    360
gtcactttat ctttacctaa aaaacaaaat gggaaagatg cagtctacac gatcaaagaa    420
gaaccaaaag acggagtgtc agctccgca aacatggttt tagctttccc tgtatatgag    480
atgatcaaac aagcagatgg ctcttataaa tacgggacag aagaactaga tactatccat    540
ctctacccta aaaatacagt cggtaatgat ggaacgttga aagttacaaa aatcggtact    600
gccgaaaacg aagcactaaa tggagcagaa tttattattt ctaaagaaga aggaacacca    660
agcgtcaaaa aatacatcca aagtgtcaca gatggattgt acacttggac aactgatcaa    720
accaaagcca aacatttcat tactggtcat tcttatgaca tcggcaacaa tgactttgcc    780
gaggcatcta ttgaaaaagg ccagttgatc gttaatcatt tagaagttgg aaaaatataat   840
ttagaagaag taaaagctcc tgataatgcg gaaatgattg aaaagcaaac aatcacgcct    900
tttgagatcc tggcaaatag ccaaacacca gtagaaaaga ccatcaaaaa tgatacgtct    960
aaagttgata aaacaacacc tcaattgaat ggaaaagatg tcgcaatcgg tgaaaaaatt   1020
caatatgaga tttctgtcaa tatccccatta ggtatcgctg ataagaagg aacgcaaaac    1080
aagtacacaa cattcaaact tatcgatact catgacgctg ctttaacatt tgataatgat    1140
tcttcaggaa cgtatgctta tgccttatat gatgaaaata aagaaatcga cccagtaaat    1200
tattctgtca ctgagcaaac agacggattc acggtttcag ttgatccgaa ttatattcct    1260
tcattaactc ctggcggtac attgaaattc gtttactata tgcatttgaa cgaaaaagca    1320
gatccaacca aaggattttc taaccaagca aatgtcgata acgggcatac aaatgatcaa    1380
acaccaccgt cagtcgatgt cgttactggg ggcaaacgat ttgttaaagt agatggtgac    1440
gttacatcag accaaacact tgctggagca gaattcgtcg ttcgtgatca agatagtgac    1500
acagcgaaat atttatcgat cgacccatcc acaaaagccg tcagctgggg atcggcgaaa    1560
gaatcagcaa cggttttttac aaccacaagt aacggtttaa tcgatgtgac aggtctaaaa    1620
tatggcacgt actatctgga agaaacgaaa gcgccagaaa aatatgttcc attaacaaac    1680
cgtgtagcat ttactatcga tgaacaatct tatgtaacag caggacagtt gatttctcct    1740
gaaaaaatac caaataaaca caaggtaca cttccttcaa caggcggtaa gggaatctat    1800
gtgtatatcg gtgcaggagt agtccttcta ctgattgctg gactgtactt tgctagacgc    1860
aagcacagtc agatttag                                                 1878
```

Seq ID 348

```
atgaataacc gaatccttga aacattggaa tttgaaaaag tcaaacagat ggtcagacaa    60
tttgtagtga ctgcacaagg aaaagaagaa ctggcagaat tggtaccagt gagtgagaaa    120
caaacgatca caaactggct tcaagaaact gaggatggat tgaaagtgca acgtctgcga    180
ggcgggatcc ctataccctaa acttgaaaat atccgtcccc atatgaaacg gatcgagatc    240
ggagcagact tgaatggagt ggagcttgca caagtttcca gagtattgtc tacgacaagc    300
gaattgaatc gtttcattga tgatctttct gacagtgaaa tcgaatttgc acgtttgtat    360
atgtgggcag atcagttagt cacgatccct gtttttgagcc gccgtttaaa ggaagcaatc    420
gatgaagacg gacgcgtgac tgatgacgct tctcctgagt taaagagcat aagacagaac    480
attcggcgaa gtgaacaagc tgttcgtgaa cagttagatg ggatcgtacgt tgggaaaaat    540
gcaagtatt taagtgatgc gatcatcacc atgcgaaatg accgctacgt tatcccagtc    600
aaacaagaat accgcggtgt atttggcggt gtagtccacg atcaaagtgc ttcaggacag    660
actttatttta ttgaaccaaa acaagtggtg gatttgaata atcgtttgcg tcaataccaa    720
atcgctgaac gaaatgagat tcaacgtatt ttaagtgaac tttctgctga acttgttccc    780
catcgtcaag aaatcatcca caatgcttat gtgattggta aaatggatct gatgaatgca    840
aaagcacgat ttggtaaaga agtaaaagcc atcgtgccgg aatcagtga agacaatcat    900
gtcgtcttga aacaagcaag acatccactg atcgatcagg aaaaagtagt ttctaatgac    960
atcacaatcg gaaaagacta tcaagccatc gtcattacag gacctaaaac aggtgggaaa    1020
acgatcacgt tgaagacttt aggtctgtta caactgatgg ggcaagcagg attaccgatt    1080
cctgcaggtg aagaaagtca aatcggtatt tttgaggaag tatttgcgga tataggagac    1140
gaacagtcga ttgagcaaag tctatcgacc ttttcttctc atatgacgaa cactgtagat    1200
atttatcaa aagtcaatga aaaagtctg gtattatttg acgagctcgg tgcaggaaca    1260
gatccgcaag aaggtgctgc tctagcaatc gcaatcttag atgatcttgg taaaaagtct    1320
gcttatgtga tggcaacaac gcattatcct gaattgaaag tttatggata caatcgggca    1380
aatacgatta atgccagcat ggagttcgat gtggatacac taagtccaac gtatcgtttg    1440
ctgattggcg ttcctggacg aagcaatgcg tttgaaatct ctagccgatt aggtcttgat    1500
acagaagtta ttgatgaagc gaagcagttg atgaatgatg agagccaaga tttaaacgaa    1560
atgatcacag atttagagaa tcgccggaaa atggcagaga ccgaatatcg ggaaatgcgt    1620
cattttgttt ctgaggcaca agaacttcac gatgatttga agaagcttta cagctatttc    1680
tttgaagaac gtgaaaaaga aatggaaaaa gcaaacaaac agcaaacga agtcgtgtca    1740
gaggcggaag aaaaagcaga aaaaatcatt gctgatattc gtaaaatgca acagcagatt    1800
gggcaaggaa atgtgaaaga gcatcagttg atcgatgcaa agaccaattt agcaaatcta    1860
catcaagaag aaactttgaa gaaaaataaa gtattgaaga aggcgaaaga acaaaaaacg    1920
```

```
ttgaaaccgg gagacgaagt gttggttacg acttacggcc aaagagggac gcttttacgt   1980
aaaaacggga atcagtggca agtagaaatc ggtatattaa aaatgaatgt gtccgaagac   2040
gagctgacgc ctgttgctcc tcaaaaagag ccgacacaac gagttatcca tgctgtacga   2100
tctgaatcta gtagtcacgt gccgaatcaa ttggaccttc gaggaaaacg atacgaagaa   2160
gcgctaagtc aggtagatca atattggat tctgccattc tagctggtta tcctcaagtg    2220
actattgtcc acggaaaagg aacaggcgca cttcgcaagg gaattactga ttatcttaaa   2280
aaccatcgga gcgtcaaaag ctttgagttt gcaccggcaa accaaggcgg aaatggtgcg   2340
acaattgtta aatttaagta a                                            2361


Seq ID 349
atgttcgatg gctactgttc acgtgtcgta aatggcaatc caaacgttaa accagcaaca     60
agaaaaaaag tcttggaagt catcgatcgc ttggattatc gtccaaacgc agttgcgcgt    120
ggcttggcaa gtaagaaaac aacaacagtt ggggtaatca tccctgatgt cagcaatatg    180
ttctttgctt ctttagcacg tggtatcgat gatgtggcaa caatgtacaa atacaatatt    240
atcttagcaa attcagatgg caacgaccaa aaagaagtca atgtattgaa taacttactg    300
gctaagcaag tggacggcgt gatcttcatg ggtcaccata tcacagacga aatccgtgga    360
gaattttcac gttcaaaaac tcctgttgta ttagctggct ctattgatcc agatgaacaa    420
gtaggtagcg taaacattga ttatgtatcc gcaacaaaag atgcagtaaa taaattggca    480
aaaagcggca ataaaaaaat cgcttttgtc agtggtgcgt tgatcgatcc tatcaatgga    540
caaaatcgtc tgaaaggcta taagaagca ttaagcgaaa ataacctacc atatagcgaa      600
ggattgatct ttgaagcaca atacaacttc aaagatggtc tatctctagc cgatcggatc    660
cataatagcg gtgcaacagc tgtctatgtt tcagatgatg aattagccat cggtattttg    720
gatggtttat tagatcgtgg cgtgaaagta ccggaaggat tcgaaatcat cacaagtaac    780
aattcactac tgacagaagt agctcgtcct cgtttaacaa gtatcacaca accattatac    840
gatttagggg cggtatctat gcgcctattg acgaaaatga tgaataaaga agaagtcgaa    900
gaaaaaacga ttatcttgcc ttatggaatc gaagaaaaag gatctacaaa ataa          954


Seq ID 350
atgaaaaaaa ttgatatttt atctcgattg aaaaacgccg gagtgattgc ggtggtacgc     60
ggaaaaagca aagaagaagc tttgaatgct tgccacgcaa tcatcaaaga tggtctaact    120
ggtattgagt tgactttttac agtccctcaa gcagatcaag taatcaaaga acttttgtca    180
ttctacaaag accaaccaga aatagtgatt ggcgcaggca cagtattaga cgctgtcaca    240
gctcgcttgg ctatcttagc tggcgcagag tatattgtta gcccttcctt tgatcaagaa    300
acggcagaga tgtgcaatct gtatcaaatt ccttacttac caggatgtat gactattacc    360
gagatcaaaa cagcgctgaa aagcggggta gatatcgtca aactatttcc tggaagcgct    420
tatggaccaa gtatcatttc tgccttcaag gcgccaatgc ctcaagtcaa tatcatgcca    480
actggcggtg tgagtctgga caatatgaaa gagtggttcg atgctggtgt cgtaacagtc    540
ggagtaggtg ggaacctttt agcaccgca gctacaggag attttgacaa agtaacggaa     600
gtcgcgcaac agtatgcagc aaaaatgaaa gaaatcaaga ggtag                    645


Seq ID 351
atggatagtg tagtggtcca tcgtaaatcg cgtttatggg aatttacgct gcattttcc       60
aaaattttgc ctattatgct atatcggtca ttgactcagc atctgactat cgcttttaaa     120
gatattgcac aagtcaaact gaatatctta gcagatgacc agacattcga tgaacagctc     180
ttgcaagatt actgggccca agctttggaa aatcagcaat gcgatacacc actagcacag     240
caagtattga aaacgcaggt acctgtgata aaagatcgta aagtcatctt gccaatcgac     300
agtacaggag ctatttctta tttgaagcaa caatacttgc cattagtaga agagttgttc     360
gtcagttacg gattcccaaa atttcgtatc gaaccagaag tggatgagca gcaagcggaa     420
cgagtattga aactttttga agaacgcaaa caagaacagg cagaagcctt tatgaagcaa     480
gcagctgaaa gtctagttgt ccacgaacaa aagaaaaagg aaagaaaaga gcaatcgcca     540
gctcttgaag gacccatcca catccagcta ggacgaaata ttccagccga tgaaccaacg    600
acaccaatga tcagtatcgt tgaagaagaa cgaagagtca cacttgaggg gtacgtattc    660
gacaaagaag tcagagaact acgttcaaaa cgcaagatct tgactttgaa aatcacagat    720
tatacttcat catttatcgt caaaaaattc tcgaacggcg aaaaagatga acaagtattt    780
gatgcgatca gtgtcggtag ctggttaaag gtacgaggaa gtattcaaga agatacgttt   840
gtccgtgatt tggtcatgaa tgcgcaagat atcatagaag tcaaacatac acctcgcaaa    900
gattatgctc ctgaaggga aaaacgagta gagcttcatg ttcatagcaa tatgagtacg    960
atggatgcaa caaacagcat ctctgatctt gtggcacaag cagggaaatg gggacaccga   1020
gctatcgcca ttacagacca cggtggtgcg caagcctttc cggaagcaca tagtgctgga   1080
aaaaaagcag gtgtgaaaat cttatacggc gtggaagcga atgtcgtaga tgatggtgtc   1140
ccaattgctt ataatgatgc acatgaagca ttgagtgaag ccacatatgt cgtatttgac   1200
gtggaaacga caggattgtc tgcagtatat gacacgatca tcgagctagc cgcagtcaaa   1260
atgtataaag gaaatgtgat cgaaagcttt gatgaattta tcgatccagg acatcctttg   1320
tcacgaacaa cggtagatct gactggtatc acggatggaa tggttcgagg atcaaagtcg   1380
gaagaagaag tattgcggat gttccttgaa ttttcaaaag atacaatttt ggtagcccac   1440
aacgcggctt ttgatatggg gttcttgaat actagctacg caaggtatgg tattcctgaa    1500
gcggccaatc cagtcattga tacgttggaa ttagctagat atctgtatcc tcaattcaaa   1560
cgtttcggat taggtgtgtt atctaaaaaa tttggtgtca gcttggaaca gcatcaccga   1620
gcaatctatg atgctgaagc aactggtcac ttagcatgga tctttgtgaa agaggcaatg   1680
gacaatcata acatgcttta tcatgatcag ctaaatgagc atatcggaga aggagattct   1740
tacaaaaggg ctcgaccttt ccatgtaacg attttagcaa aaaatcaggc aggactgaaa   1800
```

```
gatttgttca aactgatttc aatgtcgaat gtcgaatact tcgaacgcgt cccacgaata   1860
ccacgctccc aattgaaaaa gatgcgtgaa aatctattga tcggctcagc ttgtgataag   1920
ggagaaatat tcgaagccat gatgcaaaaa ggggtagaag aggcaagaaa ccgagcaaaa   1980
ttctacgatt atattgaagt tatgccaaaa gcggtctatg cgcctttgat cgagcaagaa   2040
ttggtaaaaa atgaacatga tttggaagaa attatccaaa atttggtaga aatcggaaaa   2100
agtttggata aaatcgttgt tgcaacaggt aatgtacatt atctgaacga agaagatgcc   2160
atttaccgaa aaattttgat caattctatg ggcggagcga atccgttgaa tcgtcatagt   2220
ttgccggatg tccatttccg aacaaccgat gagatgctaa ctgcatttca cttttttaggg   2280
gaagaaacag cgaaagagat cgtagtagaa aatacgaata aaatcgcaga tatctgcgaa   2340
gaagtgattc cagtcaaaga tgaactttat acaccgaaga ttccaggttc agaagatgaa   2400
atatcagaat tgagttacac aaaagcaaaa caaatgtacg gagatccgtt acctgagatc   2460
attcaaaaac gattaaaaaa agagttgaat tctattaatg gtaatggctt ctcggttatt   2520
tatctgatcg cccaaaaact agtgcataaa agtaatgaag atggctatct agtaggttca   2580
cgtggctctg ttggctctag ctttgttgcg acgatgacag gaatcactga agttaatccg   2640
ttagcccctc attattattg tccagagtgc caatattccg agtttttttga agatggtacg   2700
tatggttctg gatttgacat gccagaaaaa caatgtccaa aatgtggtgc acgcttaaac   2760
aaagacggtc atgatattcc atttgagaca ttttttaggat tccatggaga caaagtgccc   2820
gatattgatt tgaacttctc aggtgattat caagcagaag cccataacta tacgaaagta   2880
ttgtttggcg aagactatgt ttaccgtgcc ggaacgattg gtacagtcgc cgacaaaaca   2940
gcatatggct atgtaaaagg atacgaaagg gacaataatc tgcagtttcg aagtgcagaa   3000
gtcgacagat tagcaaaagg cgctactgga gtcaaacgga caaccggcca gcatccaggg   3060
gggatcatcg ttattccga ttatatggat gtctatgatt ttacgccgat ccaatatcca   3120
gcagacgatc aaaattcgga atggaaaacg actcattttg attttcactc gattcacgat   3180
aatgtattga aattagatat tctagggcat gatgacccaa ctgttattcg gatgcttcaa   3240
gatttatcag ggatcgatcc acaaacgatc cctacggatg accgggaagt gatgcggata   3300
tttgccggtc ctgaagtact tggagtaagt caggaacaaa tttattccaa aacagggaca   3360
ttaggtattc cagaattcgg cactcgtttt gttcgaggga tgctagaaga aacgcatcct   3420
acaacttttg cagaactact gcaaatttcc ggactatctc atggtacaga cgtatggtta   3480
ggaaatgcgg aagaattgat tcgacgagga gatgcgacac tggcggaagt aatcggctgt   3540
cgtgatgata ttatggttta tttgatccat gctgattaag atagcgggat ggcatttaaa   3600
atcatggaaa ctgttcgtaa aggacaatgg aacaagatcc ctgatgaatt gcgagaaacc   3660
tatctatcgg cgatgaaaga aaataatgtg cccgactggt atatcgattc ttgttcaaaa   3720
atcaaatata tgtttccgaa agcccatgct gcagcttacg tattgatggc tcttcgtgtc   3780
gcttatttca aggtttattt ccctattttg tattactgtg catactttc tgtacgtgca   3840
gacgactttg atctagtctc catgtgcaaa ggaaaagatg ctgtgaaaca agcaatgaaa   3900
gaaatcacgg acaaaggatt agatgcttcc gtaaaagaaa agaatcaact gacagtgctg   3960
gagttagcca atgaaagct ggaacgagga ttcaaatttg gaatgattga tttgtataaa   4020
tcggatgccg ttaactttgt tattgaaggt gatacgttga tcgcaccatt tagagctgtc   4080
ccaagtttgg gaaccaacgt ggcaaaacaa atcgttgaag caagaaaaga cggtccattt   4140
ttatccaaag aagatcttgc cacccgcggt aaagtttcaa aaacattgat cgaatatatg   4200
aacgacaatg gcgtattgaa agatctacca gatgaaaatc aattatcgtt attcgatatg   4260
ctataa                                                               4266
```

Seq ID 352
```
atggttttat gggtaattaa tatcattttg ctattgattg tgcttgcaat catcctttgg     60
gaagtctatc tacgagtaat ggccaaacgt tcagcaacga ctttgacaga agaagaattt    120
cgagaaggaa tgagaaaagc acaagtaatc gatgttcgcg aaaaagatgt gtttgatgct    180
gggcatattt taggtgcaag aaacatccct tacactgttt tgaaagattc attaggatct    240
attcgaaaag accagcctgt ctatatttac gatcaaaaga aagtttgag catccgtaca    300
gcaaataaat tacgtaaagc aggttatcaa gagatttatt tacttaaggg tggatatgac    360
ggatggtctg gaaaaatcaa aagtaaaaaa gtataa                               396
```

Seq ID 353
```
atgataaaaa ttactttttcc agatggtgca gtcaaagaat ttgagtctgg aacaacaaca     60
ttggcaattg ctgaaagcat ttcaaaaagc ttagcaaaaa aagctttagc tggtaaagtg    120
aacggaaaac tcgttgattt gactcgtcca atcgaggaag atgcaagtat tgagatcatc    180
acaccagacc atgaagatgc acttggctta gtccgtcatt cagcagctca tttgatggcc    240
caagccatgc gtcgcctcta tccaaatatc cacttggcg ttggaccagc tatcgattca    300
gggttttact acgatacaga taatggacaa aaccaagtaa ctgcagaaga tttgccggca    360
atcgaagcag aaatgatgaa aatcgtcaaa gaaaatctgc cgatcgaacg ccgtgttctt    420
tcaaacagg aagcattaga aattttttgca agtgatccgt acaaagtaga actgattagt    480
gaattgccag aagaagaagt gattactgct tatcaacaag gcgaattcat cgatttgtgt    540
cgtggaccac atgttccatc cactggccgt atacaagtat ttaaattatt atcagtggca    600
ggtgcttatt ggcgtggaaa ctcaaacaat caaatgatgc aacgtgtata cgggacagcc    660
ttttttgaca agaaagcttt gaaagaatat atccgtctcc gtgaagaagc aaaagaaaga    720
gaccatcgta aactaggaaa agaacttgac ttgttcatgg tttccccaga agtaggttct    780
ggtttgccat tctggcttcc aaaagggca acaatccgac gcacgatcga acgttatatc    840
gtagacaaag aagttagtct aggttaccaa catgtatata caccaataat gggagatgtt    900
gagttgtaca aaacatctgg acactgggat cattatcaag aagacatgtt cccgccaatg    960
gatatgggag atggcgaaat gcttgttctt cgcccaatga actgccctca tcacatgatg   1020
gtctacaaaa acacgatcca ttcttatcgt gaattaccaa ttcgtatcgc tgaattagga   1080
```

```
atgatgcatc gctatgaaaa atcaggtgct ttgtcaggtt tacaacgtgt acgagaaatg   1140
acgttgaatg atggtcatac atttgttcgt cctgatcaaa tcaaagatga atttaaacgt   1200
acgcttgaat tgatggtggc agtatatgct gacttcaata ttacagatta tcgtttccgc   1260
ttgagctatc gcgatccaaa caatacagat aaatattttg atgatgacgc aatgtgggaa   1320
aaagcccaag caatgctgaa agccgctatg gatgaattag aactagatta cttttgaagca   1380
gaaggcgagg cagccttta cggaccaaaa ctagatgttc aagtcaaaac agctcttggt   1440
acagaagaaa cattatcaac tatacaatta gacttcttat tgccagagcg ctttgatttg   1500
acatatgttg gagaagacgg ggaaaatacg catcgtccag tagttatcca ccgcgcgaatt   1560
gtttctacaa tggaacgttt tgtcgcttat ttaacagaag tatacaaagg cgctttttcca   1620
acatggttgg cacctataca agcaacaatc atccctgttt ctgttgatgc acatggcgat   1680
tatgcttatg aaatcaaaga acgtttgcaa atgaagtggt tgcgtgtgtaga agttgatgat   1740
cgtaatgaaa aaatgggcta taagatacgt gcctctcaaa cacaaaagat cccttaccaa   1800
ttagttgttg gagacaagga attggaagat gcaacagtga acgtacgtcg ttatggaagt   1860
aaagaaacag ctgtagaaga tttgaatatc tttattgatg ctatggaagc agaagtcaaa   1920
aattacagca gagaaaacta a                                             1941
```

Seq ID 354
```
atgactacta attttggaa agagctgcca aagcctttt ttgtttttagc accaatggaa     60
gatgtaacag atgttgtatt tcgccatgtc gtaaagcatg cagcagcacc ggatgtcttc    120
tttactgagt ttacaaactc agacagtttc tgtcatccag atggaaaaga cagtgtacga    180
ggaagattaa catttacaga agatgaacag ccgatggttg cccatatttg gggagacaag    240
ccagaatttt ttagagaaat gagcattgct atggcagaaa tgggcttcca aggaatcgat    300
atcaacatgg gctgcccagt ccctaatgta gcagaacgtg gaaaaggcag tggcttgatc    360
cttcgtccag aagtagcggc acaattgatt gaagcagcta aagctggcgg tttgcctgtc    420
agcgtcaaga cgcgtatcgg ctatacggaa atgtctgaga tggaagaatg gatcagtcac    480
ctacttcatc aagatatagc caatctttct gtccatttgc gtacgcgtaa agagatgagt    540
aaagtagatg cacattggga agtcattcct aaaatcattg ctttacgtga ccaaatcgct    600
ccacaaacac taatcacgat caatggagat atccccgatc gccagacagg agaagaattg    660
gccgagaaat atggcgttga tggcatcatg attggccgag ggattttcaa gaatccttat    720
gcatttgaaa aagacccgag ggaacataca gaaaaagaac tgatcggact acttcgtctg    780
caactagatc tgcaagatca ttattcagaa atcatcccgc gttccatcac aggactacat    840
cgtttctttta aaatctatat caaaggattt ccaggagcaa atgacttgcg agttcagttg    900
atgaatacga aatctactga tgaagtacga gagatcttag agacctttga aaaagaacat    960
ccaacgcttt tcagtgaaca atga                                          984
```

Seq ID 355
```
gtggtgaaaa agtcattcaa accgtttggg gagtgggcta taaaattgac gcaagatagt     60
aaacgagcaa ctgaaaaacg ccgtcggatc aatttgactt caaaagaaat cagcgaattg    120
ttagcggaag gaatcatcac gattatcctt ttgcttctat taaatgtttc tattctggtt    180
gttatcagtt cggtgatcaa cagttcacct tctttgacga atgcgatttg ggattcaaag    240
aatattttttg cagaacgatt gaacactgat ttgttctgga acggccggaa ttttattatt    300
ccctttttct ttctattgga tattggtgtt ttgtattggc gcttgattgc aagatatcgc    360
caaatgcaat tgcggccatat catcagcgaa cttcactaca tcgcaaacgg aaaattatgat    420
catcgtatcc cttttgaact gagtggtgac ttaagccgag tggtaacaag tatcaatgga    480
ttggtggata gtaccgtagc agcgattgaa gacgaacgta aaatcgaaaa atcaaaagat    540
gaattgatca cgaatgtcag tcatgatatc cgtaccccac tcacgtccat cattggatac    600
ctaggactga tagaagatgg ccagtatcac tcagaagaag atcttttaaa gtatacgcat    660
actgcttata tcaaagcaaa gcagatgaaa tctttagtgg acgacttatt tgaatacact    720
aaggtacgcc agcctgctgt cccggtgaat ttttctgcct ttgacatgat ccagttgatc    780
gaacagttag cagctgattt tgagttggaa gcttctaaaa aaaatatcca aatacttgtt    840
caatcaaaag tagactccct tatcatggat ggcgataccg aaaaacttgt acgtgtcttc    900
aataatcttt tgacgaatgc gttgaaatac ggaaaaggag caacgaagat cgtaatcgaa    960
gtagaacgta tcggttcaga agttgttgca acagtcaaaa ataacggagc aatgatccca   1020
caacaggcaa tcgataatct attcgaccgt ttttatcgtg tagaagaatc tcgttcacaa   1080
gcaactggag gaactggact tggtttagca atcgctcaaa gcattgtggc tttgcacggt   1140
ggatacattg atgcaaaatc agataagcaa tggacttctt tcatcatgca ccttccaatc   1200
aagaaaaatg aaaaattacc tattgataca caggaagtta ttgacgaaag ctga         1254
```

Seq ID 356
```
atgaatcatt ttaaaggcaa acaattcaaa aaagacgtca ttattgtcgc tgttggttac     60
tacctgcgtt acaatctaag ctatcgtgaa gttcaggaat tgttatatga tcgtggaata    120
aatgtttgtc atactacgat ttatcgttgg gtacaagagt acagcaaagt cctctattat    180
ctttggaaga agaaaaatag acaatccttc tattcatgga aaatggacga aacctatatc    240
aaaattaagg gacgttggca ttatctttat cgtgcaattg atgcggacgg cttaacctta    300
gatatctggt tacgaaagaa acgggatacg caagcagcct atgatttctt aaaaacgactc    360
cataaacagt ttggtgagcc gaaagcaatt gtgaccgata aagcaccttc tcttggctcc    420
gcctttagaa agttacagag tgtgggttta tatactaaga cagacaccg aactgtgaag    480
tatcttaaca atttaataga acaagaccat cgacctatta aacgacgaaa taaattttat    540
caaagtctcc gtacagcctc ttccacgatt aagggcatgg agaccattcg aggaatatat    600
aaaaagaacc gaagaaatgg aacgctcttc ggcttttcgg tgtctactga aatcaaggta    660
ttaatgggaa taacagccta a                                             681
```

Seq ID 357
MVGIILASHGQFAEGILQSGSMIFGEQENVKAVILKPSEGPDDLRAKLEEAVASFDNQDEVLFLVDLWGGTPFNQSNTLFE
EHKDKWAIVSGLNLPMLIEAYASRFSMESAHEIAAHIIETAKDGVKVKPEELEPARAPKAAVEDAQPKGALPEGTVVGDGK
IKYVLARVDSRLLHGQVATANTKAVQPNRIIVVSDAVSKDDLRKRLIEQAAPPGVKANVIPISKMIEVAKDPRFGNTKALL
LFENPEDVLTAVEGGVDIKELNVGSMAHSVGKVVVSKVLSMGQEDVEAFEKLEQKGVKFDVRKVPNDSRDNMDDILKKAKA
ELAKA

Seq ID 358
MIIKEEIFVPKELVKEKIDLLVENLTTIKDDNGEFLLDFDGLKVDDKSWTVWNWPQGVGLYGIYKNYQMTKNPRAYQVVNE
WFEDRMEEGAPPKNVNTMAPLLTMAYLYEDTKDSRYLPYLEQWAEWVMNDMPRTNEDGLQHATYGPENKNQLWDDTLMMTV
LPLAKIGKLLNRPEYLEEARYQFMIHIKYLMDKRTGLWYHGWTFEGNHNYAEAFWARGNCWLTIAIPEIIEILELSKEDAL
RKLLIETLEAQVRALKTYQSESGLWHTLLDDPSSYVESSATAGFAYGILKAVHKRYLPQEYKEVAYRAIQGLLEQIDEKGE
VQNVSIGTGMGDSLDFYRNIGITAMPYGQSLTVLCLTELLVSYC

Seq ID 359
MKIAIAGAGAMGSRFGLMLHQSGNEVLLIDGWAEHVQQIKEHGLQANFNGKEVEAKLPIVLQSEVEKEDQVDLIILFTKAM
QLEKMLQDIQSLIKKDTEVLCLLNGIGHEDIIEKFVPMENIYIGNTMWTAGLEGPGQVKLFGSGSVELQNLGDGKEAAAKK
LADKLSESGLNAHFSDNIHYSIYRKACVNGTMNGLCTILDVNMAELGKTSTAHKMVATIVNEPAKVAAVEKIELDVPEVIA
HCESCFDPETIGLHYPSMYQDLIKNHRLTEIDYINGAISRKGKKYGVATPYCDFLTELVHAKEDSLNVK

Seq ID 360
MTFPYSEGNKRYHSWNYALRNEFGGKIFKVPIDGGFDCPNRDGTVAHGGCTFCSVSGSGDMIVAPEDPLPIQFQKEIDMMH
QKWPHVQQYIVYFQNFTNTHAPVEIIKERFEQVVNLPGVVGLSVGTRPDCLPDEVVDYLAELNERMYLWVELGLQTTFEET
SKLINRAHDYQTYLDGVAKLRKHNIRVCTHLINGLPGESLEMMKENVRRTILDSDIQGIKLHLMHLMRKTRMLRDYHEGRL
QLMSRPDYVNVICDQLEMIPKEIIIHRLTGDAPWDSLIGPMWSLKKWEVLNAIDEELLRRDSFQGKYDVRKKVSV

Seq ID 361
MAKTGMYVGLDIGTTSVKVVVAEYIDSQMNIIGVGNAKSEGINRGIIVDIDKTVQAIQRAVRQAEEKAGIQIKGVSVGLPA
NMLEVENCQGMIAVNGDSKEITDEDVRNVASAALVRSIPPERQIVSILPQDFTVDGFEGIKDPRGMIGVRLEMYGLLFTGP
KTIIHNIRKCVENAGLIVNEMVITPLALTESILSDGEKDPGTIVIDMGGGQTTTAVMHDKQLKFTNLDQEGGEFVTKDISI
VLNTSFNNAEALKINYGDAYPERTSPDEEFPVDVIGQSEPVKVDERYLSEIISARMEQIFNKAKDALDQIEALELPGGVVL
TGGAASLPGVVDLAQEIPGVNVKLYVPNHMGLRNPVFTNVISIVDYSANLSEVYQLAKIAVTGETSAARQMVVEQETTNTY
ESYEAPEETIYDEPEPKESGENVKNKIKGFFTNIFD

Seq ID 362
MAKKTIMLVCSAGMSTSLLVTKMQKAAEEKGMEADIFAVSASDADNNLESKNVDVLLLGPQVRFMKAQFEQKLAPKGIPLD
VINMQDYGMMNGEKVLAQAENLMK

Seq ID 363
MKNHKKINVMLGVLFLILPLLTNSFGAKKVPAEETAAQVILHKKKMTDLPDPLIQNSGKEMSEFDQYQGLADISFSVYNVT
QEFYAQRDKGASVDAAKQAVQSLTPGTPVASGTTDADGNVTLSLPKKQNGKDAVYTIKEEPKDGVSAAANMVLAFPVVYEMI
KQADGSYKYGTEELDTIHLYPKNTVGNDGTLKVTKIGTAENEALNGAEFIISKEEGTPSVKKYIQSVTDGLYTWTTDQTKA
KHFITGHSYDIGNNDFAEASIEKGQLIVNHLEVGKYNLEEVKAPDNAEMIEKQTITPFEILANSQTPVEKTIKNDTSKVDK
TTPQLNGKDVAIGEKIQYEISVNIPLGIADKEGTQNKYTTFKLIDTHDAALTFDNDSSGTYAYALYDGNKEIDPVNYSVTE
QTDGFTVSVDPNYIPSLTPGGTLKFVYYMHLNEKADPTKGFSNQANVDNGHTNDQTPPSVDVVTGGKRFVKVDGDVTSDQT
LAGAEFVVRDQDSDTAKYLSIDPSTKAVSWVSAKESATVFTTTSNGLIDVTGLKYGTYYLEETKAPEKYVVPLTNRVAFTID
EQSYVTAGQLISPEKIPNKHKGTLPSTGGKGIYVYIGAGVVLLLIAGLYFARRKHSQI

Seq ID 364
MNNRILETLEFEKVKQMVRQFVVTAQGKEELAELVPVSEKQTITNWLQETEDGLKVQRLRGGIPIPKLENIRPHMKRIEIG
ADLNGVELAQVSRVLSTTSELNRFIDDLSDSEIEFARLYMWADQLVTIPVLSRRLKEAIDEDGRVTDDASPELKSIRQNIR
RSEQAVREQLDGIVRGKNAKYLSDAIITMRNDRYVIPVKQEYRGVFGGVVHDQSASGQTLFIEPKQVVDLNNRLRQYQIAE
RNEIQRILSELSAELVPHRQEIIHNAYVIGKMDLMNAKARFGKVKAIVPGISEDNHVVLKQARHPLIDQEKVVSNDITIG
KDYQAIVITGPNTGGKTITLKTLGLLQLMGQAGLPIPAGEESQIGIFEEVFADIGDEQSIEQSLSTFSSHMTNTVDILSKV
NEKSLVLFDELGAGTDPQEGAALAIAILDDLGKKSAYVMATTHYPELKVYGYNRANTINASMEFDVDTLSPTYRLLIGVPG
RSNAFEISSRLGLDTEVIDEAKQLMNDESQDLNEMITDLENRRKMAETEYLEMRHFVSEAQELHDDLKEAYSYFFEEREKE
MEKAKKANEVVSEAEEKAEKIIADIRKMQQQIGQGNVKEHQLIDAKTQLANLHQEETLKKNKVLKKAKEQKTLKPGDEVL
VTTYGQRGTLLRKNGNQWQVEIGILKMNVSEDELTPVAPQKEPTQRVIHAVRSESSSHVPNQLDLRGKRYEEALSEVDQYL
DSAILAGYPQVTIVHGKGTGALRKGITDYLKNHRSVKSFEFAPANQGGNGATIVKFK

Seq ID 365
MFDGYCSRVVNGNPNVKPATRKKVLEVIDRLDYRPNAVARGLASKKTTTVGVIIPDVSNMPFASLARGIDDVATMYKYNII
LANSDGNDQKEVNVLNNLLAKQVDGVIFMGHHITDEIRGEFSRSKTFVVLAGSIDPDEQVGSVNIDYVSATKDAVNKLAKS
GNKKIAFVSGALIDPINGQNRLKGYKEALSENNLPYSEGLIFEAQYNFKDGLSLADRIHNSGATAVYVSDDELAIGILDGL
LDRGVKVPEGFEIITSNNSLLTEVARPRLTSITQPLYDLGAVSMRLLTKMMNKEEVEEKTIILPYGIEEKGSTK

Seq ID 366
MKKIDILSRLKNAGVIAVVRGKSKEEALNACHAIIKGGLTGIELTFTVPQADQVIKELLSFYKDQPEIVIGAGTVLDAVTA
RLAILAGAEYIVSPSFDQETAEMCNLYQIPYLPGCMTITEIKTALKSGVDIVKLFPGSAYGPSIISAFKAPMPQVNIMPTG

GVSLDNMKEWFDAGVVTVGVGGNLLAPAATGDFDKVTEVAQQYAAKMKEIKR

Seq ID 367
MDSVVVERKSRLWEFTLHFSKILPIMLYRSLTQHLTIAFKDIAQVKLNILADDQTFDEQLLQDYWAQALENQQCDTPLAQQ
VLKTQVPVIKDRKVILPIDSTGAISYLKQQYLPLVEELFVSYGFPKFRIEPEVDEQQAERVLKLFEERKQEQAEAFMKQAA
ESLVVHEQKKKERKEQSPALEGPIHIQLGRNIPADEPTTPMISIVEEERRVTLEGYVFDKEVRELRSKRKILTLKITDYTS
SFIVKKFSNGEKDEQVFDAISVGSWLKVRGSIQEDTFVRDLVMNAQDIIEVKHTPRKDYAPEGEKRVELHVHSNMSTMDAT
NSISDLVAQAGKWGHRAIAITDHGGAQAFPEAHSAGKKAGVKILYGVEANVVDDGVPIAYNDAHEALSEATYVVFDVETTG
LSAVYDTIIELAAVKMYKGNVIESFDEFIDPGHPLSRTTVDLTGITDGMVRGSKSEEEVLRMFLEFSKDTILVAHNAAFDM
GPLNTSYARYGIPEAANPVIDTLELARYLYPQFKRPGLGVLSKKFGVSLEQHHRAIYDAEATGHLAWIPVKEAMDNHNMLY
HDQLNEHIGEGDSYKRARPFHVTILAKNQAGLKDLFKLISMSNVEYFERVPRIPRSQLKKMRENLLIGSACDKGEIFEAMM
QKGVEEARNRAKFYDYIEVMPKAVYAPLIEQELVKNEHDLEEIIQNLVEIGKSLDKIVVATGNVHYLNEEDAIYRKILINS
MGGANPLNRHSLPDVHFRTTDEMLTAFHFLGEETAKEIVVENTNKIADICEEVIPVKDELYTPKIPGSEDEISELSYTKAK
QMYGDPLPEIIQKRLKKELNSINGNGFSVIYLIAQKLVHKSNEDGYLVGSRGSVGSSFVATMTGITEVNPLAPHYYCPECQ
YSEFFEDGTYGSGFDMPEKQCPKCGARLNKDGHDIPFETFLGFEHGDKVPDIDLNFSGDYQAEAHNYTKVLFGEDYVYRAGT
IGTVADKTAYGYVKGYERDNNLQFRSAEVDRLAKGATGVKRTTGQHPGGIIVIPDYMDVYDFTPIQYPADDQNSEWKTTHF
DPFHSIHDNVLKLDILGHDDPTVIRMLQDLSGIDPQTIPTDDPEVMRIFAGPEVLGVSQEQIYSKTGTLGIPEFGTRFVRGM
LEETHPTTFAELLQISGLSHGTDVWLGNAEELIRRGDATLAEVIGCRDDIMVYLIHAGLDSGMAFKIMETVRKGQWNKIPD
ELRETYLSAMKENNVPDWYIDSCSKIKYMFPKAHAAAYVLMALRVAYFKVYFPILYYCAYFSVRADDFDLVSMCKGKDAVK
QAMKEITDKGLDASVKEKNQLTVLELANEMLERGFKFGMIDLYKSDAVNPVIEGDTLIAPFRAVPSLGTNVAKQIVEARKD
GPFLSKEDLATRGKVSKTLIEYMNDNGVLKDLPDENQLSLFDML

Seq ID 368
MVLWVINIILLLIVLAIILWEVYLRVMAKRSATTLTEEEFREGMRKAQVIDVREKDVFDAGHILGARNIPYTVLKDSLGSI
RKDQPVYIYDQKKSLSIRTANKLRKAGYQEIYLLKGGYDGWSGKIKSKKV

Seq ID 369
MIKITFPDGAVKEFESGTTTLAIAESISKSLAKKALAGKVNGKLVDLTRPIEEDASIEIITPDHEDALGLVRHSAAHLMAQ
AMRRLYPNIHFGVGPAIDSGFYYDTDNGQNQVTAEDLPAIEAEMMKIVKENLPIERRVLSKQEALEIFASDPYKVELISEL
PEEEVITAYQQGEFIDLCRGPEHVPSTGRIQVFKLLSVAGAYWRGNSNNQMMQRVYGTAFFDKKALKEYIRLREEAKERDHR
KLGKELDLFMVSPEVGSGLPFWLPKGATIRRTIERYIVDKEVSLGYQHVYTPIMGDVELYKTSGHWDHYQEDMFPPMDMGD
GEMLVLRPMNCPHHMMVYKNTIHSYRELPIRIAELGMMHRYEKSGALSGLQRVREMTLNDGHTFVRPDQIKDEFKRTLELM
VAVYADFNITDYRFRLSYRDPNNTDKYFDDDAMWEKAQAMLKAAMDELELDYFEAEGEAAFYGPKLDVQVKTALGTEETLS
TIQLDFLLPERFDLTYVGEDGENTHRPVVIHRGIVSTMERFVAYLTEVYKGAFPTWLAPIQATIIPVSVDAHGDYAYEIKE
RLQMKGLRVEVDDRNEKMGYKIRASQTQKIPYQLVVGDKELEDATVNVRRYGSKETAVEDLNIFIDAMEAEVKNYSREN

Seq ID 370
MTTNFWKELPKPFFVLAPMEDVTDVVFRHVVKHAAAPDVFFTEFTNSDSFCHPDGKDSVRGRLTFTEDEQPMVAHIWGDKP
EFFREMSIAMAEMGFQGIDINMGCPVPNVAERGKGSGLILRPEVAAQLIEAAKAGGLPVSVKTRIGYTEMSEMEEWISHLL
HQDIANLSVHLRTRKEMSKVDAHWEVIPKIIALRDQIAPQTLITINGDIPDRQTGEELAEKYGVDGIMIGRGIFKNPYAFE
KDPREHTEKELIGLLRLQLDLQDHYSEIIPRSITGLHRFFKIYIKGFPGANDLRVQLMNTKSTDEVREILETFEKEHPTLF
SEQ

Seq ID 371
VVKKSFKPFGEWAIKLTQDSKRATEKRRRINLTSKEISELLAEGIITIILLLLLNVSILVVISSVINSSPSLTNAIWDSKN
IPAERLNTDLFWNGRNFIIPFFFLLDIGVLYWRLIRRYRQMQLRHIISELHYIANGNYDHRIPFELSGDLSRVVTSINGLV
DSTVAAIEDERKIEKSKDELITNVSHDIRTPLTSIIGYLGLIEDGQYHSEEDLLKYTETAYIKAKQMKSLVDDLFEYTKVR
QPAVPVNFSAFDMIQLIEQLAADFELEASKKNIQILVQSKVDSLIMDGDTEKLVRVFNNLLTNALKYGKGATKIVIEVERI
GSEVVATVKNNGAMIPQQAIDNLFDRFYRVEESRSQATGGTGLGLAIAQSIVALHGGYIYAKSDKDQWTSFIMHLPIKKNEK
LPIDTQEVIDES

Seq ID 372
MNHFKGKQFKKDVIIVAVGYYLRYNLSYREVQELLYDRGINVCHTTIYRWVQEYSKVLYYLWKKKNRQSFYSWKMDETYIK
IKGRWHYLYRAIDADGLTLDIWLRKKRDTQAAYDFLKRLHKQFGEPKAIVTDKAPSLGSAFRKLQSVGLYTKTEHRTVKYL
NNLIEQDHRPIKRRNKFYQSLRTASSTIKGMETIRGIYKKNRRNGTLFGFSVSTEIKVLMGITA

EF0008
Seq ID 373
ATGATTAATAACGTTGTATTAGTCGGAAGATTGACAAAAGATCCTGACTTACGTTACACCGCTAGTGGTTCAGCTGTTGCG
ACCTTTACACTTGCTGTAAACCGTAACTTTACGAATCAAAATGGTGATCGTGAAGCAGACTTTATCAACTGTGTGATTTGG
CGTAAACCCGCAGAAACAATGGCTAATTATGCCCGTAAAGGTACATTATTAGGTGTTCGGGAAGAATTCAAACTCGTAAC
TACGAGAACCAACAAGGTCAACGTGTCTACGTAACTGAAGTGGTTTGTGAAACTTCCAATTGTTAGAATCTCGTTCTGCT
TCAGAACAAAGAGGAACTGGCGGCGGTAGCTTTAATAACAACGAAAATGGTTATCAATCACAAAATCGTAGCTTTGGTAAT
AACAATGCCAGTTCTGGATTTAATAACAACAACAATAGTTTCAATCCATCATCTTCTCAGTCGCAAAACAATAACGGTATG
CCTGATTTCGATAAAGATTCTGATCCATTTGGTGGCTCAGGTTCATCTATCGACATTTCAGATGATGATTTACCATTC

EF0028
Seq ID 374
ATGGCAGTAAAGAAAAAGAAAGCGACTTTCTGGGAGTTTTTCCAAGGATTAGGTAAAACGTTTATGTTGCCAGTGGCTTTA
TTAGCTTTTTATGGGGATTTTGTTGGGATTAGGAAGTTCTTTTTCCAGTGAATCCATGATTGAAACGATCCCGTTTTTAGGA
AAGCCCGCAGTGAAAATTATTTTTCCAATTTATGTCGACAATTGGTGGTTTTGCGTTTGCCTATTTACCTGTCATGTTTGCA

ATGGCCATTCCGTTGGGGTTAGTTCGGAAAGAAAAAGGAATTGCTGCGTTTTCAGGTTTCGTAGGCTATACGGTGATGAAT
TTAGCGATTAATTTCTATTTAGTGCAGACGAATCGCCTCGTGGATCCTGAGCAGTTGCGGGAAGCCGGCCAAGGAATGGTT
TTTGGGATTCAAACGATTGAAATGGGTGTTCTTGGTGGAATTATCGCTGGGCTTATTGTCTATAAGTTGCATAATCGATTT
TATACGGTGCAACTACCAGATAGTTTTGCCTTCTTTTCTGGCGCACGGTTTGTGCCAATTATTACTTCATTAGTAATGGCA
TTCGTGGGCTTAGTAATTCCGTTAGTTTGGCCACTTTTCGCGCTAATGATTATGGCGATTGGCCAACTTATCCAACGTTCT
GGTATTTTTGGGCCTTTTCTGTTTGGTTCAGGGGAACGTCTACTGTTACCGTTTGGGTTGCATCATATTTTGGTTTCCATG
ATTCGTTTTTACTGAAGCAGGGGGTTCAGCGGTGGTGGCTGGCAAAGAAGTGTTTGGCGCGTTAAATATTTTCTATGCAGAA
TTGCAAAATAATTTACCGATTTCACCAAGCGCAACAGCTTTCTTGTCACAAGGGAAAATGCCGACCTTTATTTTTGGCTTA
CCAGCGGCGTCACTTGCGATGTATCATACTGCAGCTCCGGCGAATCGACATAAGATTAAAGGCTTACTCCTTTCGGGTGTA
ATTGCTACGGCTATCACTGGCATTACAGAACCAATTGAATTTTTATTCTTATTTATTAGTGCCATTATTATGGCTGTTTCAT
GTCATTATGACTGGTCTAGGTTTCATGGTAATGGCTTTATTAGGTGTTGTGATTGGGAACACTGATGGCGGGCTACTGGAC
TTTGTTATTTTGGCTTGTTACAAGGAACTTATACAAAATGGTGGTGGGTTCTCATTGTCGGCGCTATCTGGTTTGTTGTT
TATTATTTCGTCTTTAAAACAGTCATTGTAACCTTTGATTTAAAAACACCTGGACGAGACAAGGTCTTAGATGAGACTGAA
TATACCGATCAAGAAGTTCAATATAAGAAACGGGTGGCTATGATGCGCCTGGGATTTTAGCGGCTTTGGGTGGGCAAGAG
AATATTCAAGCTATTGATAATTGCATTACGCGTTTACGTTTAGTGTTAGCAGATGCTAATAAAGTAGATGATGACAAATTA
AAGCAATTAGGGGCTTTAGGCGTTGTACATTTAGATGCGCAGAATGTACAAGTGATTATCGGAACCAAAGTGACAACGGTC
CGTAATCAATTAGAGATGATTTTAGGT

EF0146
Seq ID 375
ATGAAAAAGATCGCAAGTGCAGGGGTTAAGTATTTTAGTCGCAACGGGGGTAGCAGGTATTGGGGGAAATGAAGTACAGGCA
GCAGAACAAGCGCAACCAAAAACACCTGAAAACAGTTCTACAGAACAACCAGCAGTGAAAGCTACAGAAACAACGGAGCAA
GCCATTACTGAAAAAACAGCAACAAGTAACAGAGAAACAAGCAGTTGTCGATCAAAAACAACAAGTTGCTGACACTGCGAAA
AAAGAAAAAGACCACCATTGATCAATCTGTTAAAGACCAACAAGCAGTGGTCGATCAAAAACAAAGGTGCATTGGATCAAAGT
CAACAAGCAGTGACTGACCAACAAGCGGTCGTAGAGCGAAGCAAAAAAAAGTTGTGGATGAAGCAACACCTTCAGCCATTGAA
AAAGCCAAAAACCAAGTGGCTACGGATACACAAGCTGTTGATGACCAACAAAAAGTAGTGGATCAAGCTCAAGCAGACGTT
AACCAACAACAAGCAGTTGTCGAAGAAAAAGCAAAGAAACGAATGCTGCTAAAGTGCAAAATGATAAAGATCAACAAGCA
GTAACAGCTGCGAAACAAGAACAAGCCAAGCTTGAAGAATTAGCGAAAAATGCGGAAGTGGAAAAAGCAAAGGCTGAAAAA
GAACAAGCAGCAAAAGAAGCTGAACTGGCTAACAAACAGAAAGAAGAAGCCAAAGCTAAAGATCAAAAAACGAAGGACGAT
CAAGCAGTCGCAGACCAACAAACTGTTGTGACAACTGTCAAGAGAAGTAGCAGACGCAAAAGCAGATACAGCTGCGAAA
CAGGCAGACTTAACAGCGAAAGAATGCTTTGAAAGACAAGCAAACAGCAACAAAGCAAGTGCAGAATACTTTAGATAAA
TCAAAAGAAGAGCTAAAAGGACATAAAGGAATTAACTTGCCTGCTAACTTCACACCTGATTATTACAAAAAATTATCTGAA
CAAGAAAAGCAAGCAATGGAAAAGAAGCATTAGCATTAAATAAAGTTTTTCCTGAAAATCAAGCAGATGTGGCAAAAGCA
ACGGAAATGATCAATGTCAAAAATCCTACCGGAAAACAAAAGCAACAAATGAGCGATTACGTTGTAGGACTTATCAATGAT
GTTCGCGAAAAGCTTGGGTTACAAAAGTTGAAGATTTCTAACCAAGCTATGAGATTTGCTTGGGATGTAGCAAAATATGAT
AATCCCAAAGAGTTTGATCATGACGTAAATGCAATTAATCGTGCAGCAAAAGAAAATGGTTTTAAAGAGTATCCAGGGCAA
AACTTTTATGAAAATCTAAGTATGGGCTATTTTGAAACGATTAATGGCACTATTTCTCAACTAGAGTTTGAAAAAGCTGCT
CGAAAAACAATTGCTGATATGCTCTTTGACGATGAAAGTTCAGCGTATTCTCATATAGATTCATTGCTAAAAGGGGACACA
ACAAACATGGCAGTTTCTATTTCAGGAGATTTAAATGATATTTCGGCAAAAATCCATATCATTAGCTATAATCAATCAAAA
TTAGTTGAAGCAAATACCTATGAAGAAGGCACTGCTCCAGTCTTTAAGAGTAAAGAAACCCTTCAAAAAGAAGTAGCAACC
AATCAAGAAAAAATTAGCTACTGCACAACAAGCAGAATCAGACGCTCAACAAGCAAAAAGTGCAAGTCAGCAAGTCTTAAAT
ACAGCCAAAACAACACAAGCACAGCAGAAAAAGAACTATCTGTTCATAAAGCGACATTGGCTAGTCTTCTTCAAGCAGTTGCG
ACTAAAAGTACGACAAATTATGAAGAAAAAGTACGACAAACTGCAACGGCAGAAAAAAGCCTTCAACAAACAAAAGATCAA
TTGGCAACAATCAATGAGTTAATTCAGAATCGAGCTGCTGTTTTAGAAAAAGCGAAAACAAACGTTTCGGAAGCACAAGCA
ATTGAACAAACGTCTGCTAAAGTGCTGAAAGAAAAACAAGCAGCCCAAAAAGCAGAAGAAAACACATTGAATAGCTTGAAG
GAAGTATTGAATTTAGCAAAAGAAAATTTAAATCAAAAACAAGTTGCATTTAAAACAAGTACACGTTCATTGTCTCGTTTA
GAAAATGCTCAACCAACATACGAAAAAGCATTAAACGAGTTAAACAAAGCAGAAGCAGCAGCGGTCCAAGCACAAGAAGCC
TATGAAAATTCTCTGAAATCATTGGAGGAACTTAAAGAACAACAAGCCGTTGCTACACTTGCTTATACACAAGCACAAGAA
GACCTTTCTAACGCGAAGTTAGAGCTACAGCGTACAGCAGTACCAGGGCGTATTAAGAGAATTAGAAGCACAACAAGCCGAACAGCAG
CGACAAGAAGCGTTGCAAGAACAAGTAGCAAAAGAACAACAACGCCTTGAACGAGAAGCAAAACAAAGCCAAACGTTAGTA
GCAAGTGCTACTTCAGCAGACAAAACACCTGGTCTCCAACAGTTATCTTTTTCTAAACAAAAAGAACAGCCAAAAGCACAA
GCATTAACACATTCAGAATCTCGTAAGACGAAACAAGTAGCAAAAGCCCCAGATTCTTTACCACATACAGGAGAAAAAAAT
AATAAATGGTTAGCTATAGCTGGTCTGATATTTGCTTTGTTAGGAGCTGCGGGTATTATAAGTTTTATTAGTAGAAACGAG
AAGAAAGTAAAAAAATATCTTTAAAATTAAA

EF0153
Seq ID 376
ATGAAAAAAATGATTATTATTGCCTTATTCAGTACAAGCCTTTTAGCAGGGGGAAGCAGTGTTTCTGCTTATGCGCAAGAA
TCAGAAGGAAATCTTGGTGAAACAACAGGGAGTGTTTTACCAGATGAACCGAATGTACCAACTGACCCAATAACGCCAAGT
GAGCCAGAGCAACCAACAGAGCCAAGTACACCAGAGCAACCATCGGAGCCGTCAACACCAACCGAACCTAGTGAGCCTTCA
AAACCGACGGATCCTTCGTTACCAGACGAACCGAGCGTACCAACAGAGCCGAACAACGCCAAGTAAGCCAGAGCAACCAACA
GAGCCAGAACGCCAAGTGTACCAGAGCAACCAACAGAGCCAAGTGTACCAGAAAAACCAGTAGAACCAAATAAACCAACC
GAGCCAGAAAGCCTGTGCCAGTTGTTCCTGAAAAACCAGTTGTACCACAACAACCAGAGCAACCAACAGATGTGGTGGTA
AAGCCAAATGGAGAAATTGCAACAGGAGAATCTACACAACAGCCAACTGTTCCAATTGAAACGAATAACCTTTCAGAAGTA
ACACATGTCCCAACTGTGACGACACCGATTGAAACAGCAAGCGGAGAAGCAATTGTCGCAGTGGATAAGGGCGTTCCTTTA
ACACAAACGGCTGATGGATTAAAACCGATTAAAAGTGAATATAAAGTATTACCAAGTGGCAATGTACAAGTGAAAAGTGCT
GACGGAAAAAATGAAAGTACTTCCTTACACTGGTGAAAAAATGGGCATAATTGGGTCAATCGCTGGTGTATGTTTGACTGTT
TTATCAGGAATCTTAATTTATAAAAAAACGTAAAGTG

EF0394
Seq ID 377
GTGAAGAAACGTCTATTTGCATCAGTATTACTATGTTCATTAACGCTATCAGCAATTGCTACCCCAAGCATCGCTTTGGCG
GACAATGTTGATAAAAAAATTGAAGAAAAAAATCAAGAAATTTCATCATTAAAAGCAAAACAAGGGGATTTAGCTTCACAA
GTATCTTCTTTAGAAGCAGAAGTATCTTCAGTATTTGATGAAAGCATGGCTTTACGTGAACAAAAGCAAACACTAAAAGCA
AAATCAGAACAATTACAACAAGAAATTACAAACTTGAATCAACGTATTGAAAAACGTAACGAAGCAATCAAAAATCAAGCA
CGTGATGTTCAAGTTAATGGACAAAGCACAACAATGCTAGATGCAGTTTTAGATGCGGACTCAGTTGCAGATGCAATCAGC
CGTGTTCAAGCTGTTTCAACAATCGTAAGTGCCAACAACGACTTAATGCAACAACAAAAAGAAGACAAACAAGCCGTTGTT
GATAAAAAAGCTGAAAACGAGAAAAAAGTGAAACAACTTGAAGCAACAGAAGCTGAATTAGAAACAAAACGTCAAGATTTA
CTTTCTAAACAATCTGAATTAAACGTAATGAAAGCTTCATTAGCATTAGAACAATCATCAGCTGAAAGTTCTAAAGCTGGC
TTAGAAAAACAAAAAGCAGCTGCTGAAGCAGAGCACGCACGCTTAGCTGCTGAACAAAAAGCTGCAGCTGAAAAAGCCAAA
CAAGCTGCTGCAAAACCAGCTAAAGCTGAAGTGAAAGCAGAAGCACCCAGTTGCCTCTTCATCAACAACAGAAGCACAAGCA
CCAGCAAGCTCAAGCTCAGCAACTGAATCAAGCACGCAACAAACAACTGAAACAACTACACCAAGTACAGATAATAGTGCA
ACAGAAAATACTGGCTCTTCTTCATCAGAACAACCAGTACAACCTACAACACCAAGCGATAATGGAAATAATGGTGGCCAA
ACTGGTGGTGGAACAGTTACACCAACACCAGAACCAACACCAGCGCCTTCTGCTGATCCAACAATCAATGCATTGAACGTT
CTACGTCAATCATTAGGTTTACGTCCAGTAGTATGGGATGCAGGTTTGGCAGCTTCTGCAACTGCTCGTGCAGCACAAGTT
GAAGCAGGTGGCATTCCAAATGATCACTGGTCTCGTGGAGATGAAGTTATCGCAATTATGTGGGCGCCAGGTAACTCAGTA
ATCATGGCGTGGTACAATGAAACAAACATGGTAACAGCTTCAGGAAGCGGTCACCGTGATTGGGAAATTAACCCAGGTATT
ACGCGTGTCGGTTTTGGTTACTCAGGTAGCACAATCGTAGGACACTCAGCC

EF0443
Seq ID 378
ATGAAATCAATCAAAACGATTTTACTTGGAACAACTTTGGCCGCAGGTTTAGGATTATTCTTAGGAACAGACGCAAACGCA
GAAAGCTTATATACAGTTAAAGCAGGAGATACTTTATCAACAATTTCTCATCAATTTGCAGGAGACAATAGCTTAATTCAA
AAAATTGCTTCTGATAACAAATTGCCAAACCTTGATTTAATTTTTTGAAGGAGAACAATTAGTTATTCGTTCAGAAAAAGAA
GTTGCTAATACTCCAGCACCAGCTGTAGAAGTTGCACCAGTTCAACAAGTAGTTGAACAACCTGTTGCACAACCAGTACAA
CAAGAAGTACAACAACCTGTTGCTCAAGAAGTAGCGCAACCAGCAGCACCTGCTGCAAGCAGTGATGCAAAAGAGTGGATC
GCACAACGTGAATCTAGTGGTTCTTACGATGCAACAAACGGTCAGTATATCGGTCGTTACCAATTATCTGCCTCTTATTTA
AATGGTGACTATTCACCTGCCAACCAAGAACGCGTGGCAGATCAGTACGTAGCAGGTCGCTATGGTTCATGGGATGCAGCT
AAATCATTCTGGTTAGCAAATGGTTGGTAC

EF0568
Seq ID 379
ATGAAAAAAATCTATCAATGGGCAGTGGGACAATCCTTTAAAAAATTGGATCCACGCCAACAAGTTAAAAATCCCGTGATG
TTTGTCGTTTATTTAGGTGCGCTCATCACTACAATATTATGCTTTATGGTTTAACATTTCA
ATAACCATTTTCTTTGGTTGACTTTGCTTTTTTGCGAATTTTGCTGAAGCTGTTGCAGAAGGACGAGGGAAAAGCACAGGCC
GATAGTTTGAAACAAGCAAAAAAGGAAGTCATGACTTATAAAATTAACAGTTTAGAAGATATCAAAGAAGAGAATTTCATT
GAATTACAGTCTTCTGATTTAAAGCGAAACGATTTGGTCTATGTACGTGCAGGAGAGCAAATCCCAGCTGATGGTGATGTA
ATCGAGGGGGCTGCATCAGTAGATGAAAGTGCCATTACTGGCGAATCAGCGCCAGTAATTCGTGAATCTGGTGGGGATCGC
AGTGCGGTTACTGGCGGAACAACCGTCGTTTCCGATTATTTAGTGATTCGTGTTACCTCGGAAAACGGTCAATCCTTCTTA
GATAAAATGATTGCCATGGTAGAGGGAACTCAACGAAAAAAGACTCCGAATGAGATTGGTTTGCAGATTTTCTTGATTACT
TTGACCATTATTTTCTTAACTGTATCAATTACTTTAGTCCCATTTACTGATTTCAGCAGTCAGTTATCTGGTAAAGGAGAA
GCTTTATCACTGGTTATTGTGATTGCTTTACTAATCTGTCTGGCGCCAACGACGATTGGGGCCTTAATCTCTTCGATCGGG
ATTGCAGGTATGAGCCGCTTGACGAAGGAAAACGTCATTGCCATGAGTGGTCGTGCGATCGAAGCAGCTGGTGACGTGGAT
GTTCTTTTACTCGACAAAACAGGAACCATAACATTGGGAAATCGACGAGCAAGCGATTTTTTGCCAGTCCATGGTGTTAGT
GAAGAACAGTTAGCGGATGCGGCGCAATTATCTTCTTTAGCTGATGAGACAGCAGAAGGTCGAAGTATCGTGATTTTAGCA
AAAGAACGTTTCAACTTGCGGGAAAGAGAATTCCAGCAATCAGAGGTTAAATTTATCGATTTCAGTGCCAAAACACGCATG
AGCGGAATTGATTATCGCGGTGACGTGATTCGCAAAGGTGCCGCAGACACCATGAAAAAAATATGTACAATCTAAAGGAGAG
GACTATCCATCTGAATGTGACAAGATTGTTGATAAAATCGCTCGAGCAGGTGGAACACCTCTAGTAGTTATAAAAAAATAAT
CGTGTGATGGGTGTCGTCTATTTGAAAGACATCGTTAAGAATGGTGTAAAGGAAAAATTTGCTGATATGCGTAAGATGGGT
ATCAAGACCATCATGATCACTGGGGATAACCCCTTAACAGCGGCAGCGATAGCAGCGGAAGCTGGAGTAGATGATTTCTTG
GCAGAAGCTACCCCAGAAAATAAAATGAATCTGATTCGCGAGTATCAAGAAAAGGGTCACTTAGTTGCAATGACTGGTGAT
GGGACAAACGATGCACCAGCATTAGCGCAAGCCGACGTTGCGATGGCGATGAATACAGGGGAAACAAGCCGCGAAGGAAGCG
GGGAACATGATTGATTTGGATTCCAGTCCAACAAAATTGCTCAAGTCGTACAGATTGGGGAAACAATTGTTGATGACTCGC
GGCGCGTTAACTACGTTTAGTATCGCAAATGATATCGCTAAGTATTTTGCGGTGATTCCAGTACTATTTTACAGTATTTAT
CCGCAACTGGATCGATTGAATATTATGGGATTAGGTAGTCCACTGACTGCGATTCTATCCGCTGTCATTTATAATGCAGTC
GTTATCGTGGCCTTGATTCCACTAGCATTAAAAGGTGTGAGGTATCAAGAAAAACCGGCCAGTCAAATTTTGAGTCATAAT
CTGCTGATTTATGGACTTGGGGGAATTATTGCACCATTTATTTTTATCAAGATTATAGATCTGATTCTCTCGCTGATCATT
TTA

EF0591
Seq ID 380
ATGTTTTTTTACTGGTAAAGAAAGGAATTTGAAATTTGCTATTAAGAAGAAACACGGTGGCGGTGGTGCAGCATCAGTAATT
ATCGGATGTATTATTTTTCGGAAGTATTGCGTTTGGTGGGACAACATTAGCTTACGCTGATGAAGTGCATAATAGTATAAAT
CAGGATATACAAGATTCTGGTAGTACAATTATTGGAGAAAATGATTCTTCTACCAAATCAGCTGAGTATAAAATGATTCAT
GAAATTGATGGAACTAAAATTAGTAACGGTGAAAATAGTAAAGAAACAACTACAAGTTCAGGAACTATACTGGCTGAAGAA
GCAATAGAAAGTTCAAATCAAAAAAAATTCAAAGACAAGTGAAGTCGAACAGGATCTTCATAAAGATGTATCAGGATCTGAA
TCAGTAAAAACAAGTAGAAACTTCTGATTCTATAAAAAAAATCTGAAGAATCAGCTGTTAAAACATTAAATCTGGATGATTCA
CAAGAGAATACTAATTCAATAACTACCAAGGCAGAAAATGATGCGCTATCTACAGTTAATGATGAAAAAGTATTAAATGAA
AGTGATAGTATTATCAAATCAATTCCTTCGGAAACAGAGAATGTCGATAATAATAGTAAAAGTGATAATAGGGACTTAGTA
GAAAACTCAGTGGAAACTACTAAGGAAATAATGAAAGAGGCTGATAAAACTGAAATAGCTAAAAGGAATGACAATATCTCT

```
GTATTGAATGCTGGAACTGAATCATTTAATGATAGTGATAAAATAAGAAGATCTAAAAGAGATAATGAATCTCCTGTTCCA
AGTTTGAGTATTCAAATGCCAGTCACAAAAGAGATTTACGAAAATGATAGTACTATTAAAGGAAAGTCTGAACCAGGAGTA
ACGATAAAAATTTATAAAAATGGTATTGAACTGGGGTCTATTAAAGCAGACGATGATGGAAATTTTGAATATCCTCTAGAA
TCACAAGCTATCAAAAGTGATAACTACTCTTTCGTTGCGGAAAAAGATGGAGCGAAAAGTATAGAGAGCTCTGTAAATGTA
GTAGAAAATAAAGAAATAGTAGGTAATCAACGAGATGCGTATACTCCATATGGGTTTGGTATATATAATCGAGGTTTAGTT
ACTACTGATGATGGAAAGAAACATTATACAGTTGATATGGAGGGGAATTTAGCATCTGAAGCGTACTTTGATCTTGCAGGC
GGAACAATGTACTATTCAATTGATGAAAAATTTGCACCATATGTTGAAAAAAATAGTTCTAGATGGTAAAACAATTATTTCT
AAAAATCCGTATGAATTGCCCAATAAAACCAATATTTGGTCTTCTGGTATTCTTAGTACAGATAAAAGAGGAGGATTGGTT
CGTGCTGCTTTAGTGGGTCGTACTAACGGAACAATAGATATTTATTTTAAAGATGATACTCCTCAAGAAGTTCTCAATAGT
GAAATACCTTTCCAAGTTTGGGCTAGATTTAAAGAGAAAAAAGAGGGAGTAAAAAATGATCTAATTTCTGATTTTGATCTA
AATATTTTACTTCAAAACAATAAATTAATATTAAGTAAAGATACCTTTTTCAAAGATGAAGATATTGTCAATAATAATGGT
CCTAAAATTAGTTCTAATTTTAACTATACCAATAATACTATTGATATTAATTATGGAGTAAATACTGGTAATTATGGCTTA
ACGATAGGGAAACGTAGTCCTTATAACTTGCATATTAATACAGGGAAAGAATTAAGTGATTTAGTTGATACGATAAAGATT
GACAATAAAAATTATGAGTTCGAAAAACTTCCTGATGGTTCGTTAGTGATTAAAGATATATACAAAAAGGGAATTTTAGGT
GGAGCATTACTTAACCGCCAAGAGATTACAATCAGTTTAGGATTAATAGCGAAAAAAAATCTGGGCGATCTCATTGATTTC
AAACAAGAATTATTAAAACTTGATGTATCTATTAGAGATAATACAACCTTTAAGCTAGTAAAAGGAGAAGCTACAGAAGCT
ATCCGTAAAATAGATTCCTCATTCGATACATTTAAACAGGAATTAACGAAATGGATTGAAAAAAGAGATAGTAATCCACAG
GAACCTATAACTATTAATTCTGGAGACAATACTTCTATAGTGGATCCTAAAGAGATGGTTGAATTTGCAGTAAAAAATCCA
ACCACAGAAAATATTAATAAAGCACGTAGTTTGGTTTCTAAAATGTGGCTTTGGGACAAAAACAAATATAATCCAGTTCTT
AATGCTGCTGAAGAAGTAAACAGATATCTTTCAACTACGTTAAATGGTTCATCAAATGATGATGGCTTGATATTAAATAAT
ACTCATGCGGCAACTTTCGTTAAATTAATTGATACGGATAAAGATGGGATACTTGATAGATACGAAGCAAACATGGGAATT
GGAACCAACTACACAAATGCTGATTCTGATGGTGATGGAAAAAGTGACGGATTTGAAATTTTTAATGATACTGATCCTTTG
GTGAGTCCGTACGATTGGTTTGATAAAAATGGAGAAAAAATAAAAATAGTAACAACCGATACTGATACGATAAGTGGAAGA
ATAGGAAATAACAACTATGATACTGAAAATGTCTACCCTAGAACTGTTCAATTAATAAAGGTTACTGATAACGGAGAAAAT
CTGATTTCTGAAATTAGTTCTTCTGAAGATAGAAAAGGGACTTTTGAATTTACTGGATTGAGTGGTAAATTAAGCAAGGGA
GATAGACTCGTTGTTAAAATTATTACTAATGAAGTACAAAGAGAAATTGATCGACAAAGAAATGTTGTTCAAGTTGGTTAT
GATAATCCAGAGGTATCTGAAGAAGTAATTGTTCAGGGCGCACAAGTAACAGTCTCTTTTGACTTTAATTATAATCAAACT
GATCCTGACTTTGACCGTCAACCAGTTACTCAAGTAGTAGAATTAGAAAAAGGTTCAAGGCTTTTAAATAGTTATTTTACT
CCAGAACGAAAAGGCTATAAATTTATTGGATGGAATACTGATAAACTTGGAGTAGGACAAACTGTCAATGATGGTTCTAGT
TTTAATGAAGATATAACGGTATTTGCTCAATGGGAAAAAGAACCTAACCTAGCTGGAAGAGCTACATGCTCCACAAGGTCCT
ATTGAAGAAGGTAAAAGTATTGTTCTAAATAATGTAGTACAGTCCAATAAATCAGGGTCTGTAATTGTAACTGGAGAACTT
CCAAAAGGGCTCTTTGGGTTATCTATAAATGAAACAGGTAATCTAATTGGTACTCCACTTATCAATGACTGGATTGATGGA
GAAAATAGTCGGGAAGTAAAGATTCCAGTTACAATTAGTAATGGTGACGAAAAGGTTATGGTGGAAGTTCCATTAACGATC
CTGCATGATACTCGAGCGGCAGACGCAGTAAAAGCAGCAGAAGACGCAGGCAAAGCTGGAGCAGATAAGAAAGCCGAAGTG
GAAACCGACGGTTTAGTGACTCCAGAGGAAAAAGCGGCAGTGGATGGCTTGAATGACACGACTACCGCGAAGAAAGAAGAC
GCAAGCAAGTTAGTGGATGCCTTGCCAGAAGGACCAGTAAAAGATAGCTTGAAAGATCGCTTAGATAAAGTGACAACCTCT
GAAGTAACAGTGAATGATGCAGATAGCAATGGCAAAGCGGACGATGTAGATTTAGCTGAAAAAGCAGCGGCAGACGCAGTA
AAAGCAGCAGAAGACGCAGGTAAAGCTGGAGCAGATAAGAAAGCCGAAGTGGAAACCGACGGTTTAGTGACTCCAGAGGAA
AAAGCGGCAGTGGATGGCTTGAATGACACGACTACCGCGAAGAAAGAAGACGCAAGCAAGTTAGTGGATGCCTTGCCAGAA
GGACCAGTAAAAGATAGCTTGAAAGATCGCTTAGATAAAGTGACAACCTCTGAAGTAACAGTGAATGATGCAGATAGCAAT
GGCAAAGCGGACGATGTAGATTTAGCTGAAAAAGCAGCGGCAGACGCAGTAAAAGCAGCAGAAGACGCAGGTAAAGCTGGA
GCAGATAAGAAAGCCGAAGTGGAAACCGACGGTTTAGTGACTCCAGAGGAAAAAGCGGCAGTGGATGGCTTGAATGACACG
ACTACCGCGAAGAAAGAAGACGCAAGCAAGTTAGTGGATGCCTTGCCAGAAGGACCAGTAAAAGATAGCTTGAAAGATCGC
TTAGATAAAGTGACAACCTCTGAAGTAACAGTGAATGATGCAGATAGCAATGGCAAAGCGGACGATGTAGATTTAGCTGAA
AAAGCAGCGGCAGACGCAGTAAAAGCAGCAGAAGACGCAGGTAAAGCTGGAGCAGATAAGAAAGCCGAAGTGGAAACCGAC
GGTTTAGTGACTCCAGAGGAAAAAGCGGCAGTGGATGGCTTGAATGACACGACTACCGCGAAGAAAGAAGACGCAAGCAAG
TTAGTGGATGCCTTGCCAGAAGGACCAGTAAAAGATAGCTTGAAAGATCGCTTAGATAAAGTGACAACCTCTGAAGTAACA
GTGAATGATGCAGATAGCAATGGCAAAGCGGACGATGTAGATTTAGCTGAAAAAGCAGCGGCAGACGCAGTAAAAGCAGCA
GAAGACGCAGGCAAAGCTGGAGCAGATAAGAAAGCCGAAGTGGAAACCGACGGTTTAGTGACTCCAGAGGAAAAAGCGGCA
GTGATGGCT
```

EF0592

Seq ID 381

```
TTGCCAGAAGGACCAGTAAAAGATAGCTTGAAAGATCGCTTAGATAAAGTGACAACCTCTGAAGTAACAGTGAATGATGCA
GATAGCAATGGCAAAGCGGACGATGTAGATTTAGCTGAAAAAGCAGCGGCAGACGCAGTAAAAGCAGCAGAAGACGCAGGY
AAAGCTGGAGCAGATAAGAAAGCCGAAGTGGAAACCGACGGTTTAGTGACTCCAGAGGAAAAAGCGGCAGTGGATGGCTTG
AATGACACGACTACCGCGAAGAAAGAAGACGCAAGCAAGTTAGTGGATGCCTTGCCAGAAGGACCAGTAAAAGATAGCTTG
AAAGATCGCTTAGATAAAGTGACAACCTCTGAAGTAACAGTGAATGATGCAGATAGCAATGGCAAAGCGGACGATGTAGAT
TTAGCTGAAAAAGCAGCGGCAGACGCAGTAAAAGCAGCAGAAGACGCAGGTAAAGCTGGAGCAGATAAGAAAGCCGAAGTG
GAAACCGACGGTTTAGTGACTCCAGAGGAAAAAGCGGCAGTGGATGGCTTGAATGACACGACTACCGCGAAGAAAGAAGAC
GCAAGCAAGTTAGTGGATGCCTTGCCAGAAGGACCAGTAAAAGATAGCTTGAAAGATCGCTTAGATAAAGTGACAACCTCT
GAAGTAACAGTGAATGATGCAGATAGCAATGGCAAAGCGGACGATGTAGATTTAGCTGAAAAAGCAGCGGCAGACGCAGTA
AAAGCAGCAGAAGACGCAGGCAAAGCTGGAGCAGATAAGAAAGCCGAAGTGGAAACCGACGGTTTAGTGACTCCAGAGGAA
AAAGCGGCAGTGGATGGCTTGAATGACACGACTACCGCGAAGAAAGAAGACGCAAGCAAGTTAGTGGATGCCTTGCCAGAA
GGACCAGTAAAAGATAGCTTGAAAGATCGCTTAGATAAAGTGACAACCTCTGAAGTAACAGTGAATGATGCAGATAGCAAT
GGCAAAGCGGACGATGTAGATTTAGCTGAAAAAGCAGCGGCAGACGCAGTAAAAGCAGCAGAAGACGCAGGCAAAGCTGGA
GCAGATAAGAAAGCCGAAGTGGAAACCGACGGTTTAGTGACTCCAGAGGAAAAAGCGGCAGTGGATGGCTTGAATGACACG
ACTACCGCGAAGAAAGAAGACGCAAGCAAGTTAGTGGATGCCTTGCCAGAAGGACCAGTAAAAGATAGCTTGAAAGATCGC
TTAGATAAAGTGACAACCTCTGAAGTAACAGTGAATGATGCAGATAGCAATGGCAAAGCGGACGATGTAGATTTAGCTGAA
AAAGCAGCGGCAGACGCAGTAAAAGCAGCAGAAGACGCAGGTAAAGCTGGAGCAGATAAGAAAGCCGAAGTGGAAACCGAC
GGTTTAGTGACTCCAGAGGAAAAAGCGGCAGTGGATGGCTTGTTAGAAATAAAACAGTCTTCATTTATGCCGTTTGAAAAT
```

TTATTTTCGACTACAAATGATTACTCACAGTTTCCTAAAACTGGTGAAAAATCTGATTCTATTTTAACCATTTATGGAGGT
TTATTATTCTTAAGTAGTATAGGATTATTAGGAATAAAAAAAAGAAAAAATAATACGAAT

EF0658
Seq ID 382
ATGAAGCGAAAAATTAATTATACACATATTGCTGAACCTACTGTTTTTGAGAATGTTTTTCCAACGTATTTGACGAATGAA
ACGATGATGGCGCGCAAGCAGAAAGTATTACAACGAATGGAAACAGAAAAATTCGATCAATTGGTCTTCTATGCAGACAAA
GAGCATGGCAGTAATTTTGAATACTTAACTGGCTTTATTCCGAGGTTTGAAGAAGGACTTCTTGTGTTAAATAAGGATGGT
GCGGCGACACTTATTTTAGGTAATGAAAATTTAAAACTGTGCCAACATGCCCGCATCTCAGCCGATTTAATTCATTATCTA
GCTTTTTCACTACCTAATCAACCAATGGCTGGGGAACAAAAGTTGTCGCAGATTTTTGAAACGCTTCTCGATGAAACTGCT
CAAAAAATAGGTATCGTTGGGTGGAAAATGTTTACGACACAACAGCAAGAACCTGCTACTACTATGTTTGATGTGCCACATTTT
ATTGTCGAAGCACTGAAAAAAGCGCTGCCGAAAGAAGCTCGTTTACTAAATGGGACCCACTTATTGATTGGAGCAAAGGGG
GCTCGCGCAACAAATAATGCTAACGAATTGGCGCACTATGAATATGGGGCGAATTTGGCTTCTCGCAGTATGTTAAAGGCC
TTAAATGCAATTGAAATAGGACAGCGTGAAACGGACATTGGGGCTTTGTTAAATGATGAAGGTCAAACGCCGACGGTAGTA
ACCATTGCGGCAACTGGTCAACGTTTTGAATATGCAAACATGTATCCGACTGCCAAAGAAATTCAGTTAGGTGATGCCTTA
TCATTAACTACAGGCTACAAAGGCGGCTTGTCTAGCCGAACAGGTTTTGTTATTGAAAACGAGCAGCAATTGCCAGAAGCG
CAACGTGATTATTTAGAGCGCGTAGCAAAACCCTATTTTCAAGCTGTCGTTCACTGGCTTGAAACCATTCGAATCGGACTG
TTAGGGCGAGAAATGTACCAAGCGATTGAAGAACAGTTACCAAAAGAAATCTATCACTGGCATTTGAATCCTGGGCATTTA
GTTTCGGATGATGAATGGATGTCGTCGCCTATTTATCCAGATTCAGCCATTCGCTTGGAAAGTGGCATGCTCTTTCAAGTG
GATATTATTCCTTCTGTACCTGGTTATACAGGGGTGAGTGCAGAAGAATGTGTGGCTTTGGCTGATGAAACACTGCAAAAA
GAGATTCAGCAAACTTATCCAGACATGTGGCAACGGATTGCGACACGCAAAGCCTATTTAAAGGAGACATTAAAGCTCGAT
TTACCATCGGAGGTTTTACCAATGTCTAATTTAGTTGGTTACTTACGGCCTTTTTATTTAGCTAAAGACAAAGCTTTGTGT
GTAGAAAAACCAGCGCTTAAA

EF0727
Seq ID 383
ATGAAAAAAGCTAGAGTGATTTATAATCCAACGTCAGGAAAAGAGTTAATCAAAAAGAACTTAGCCGATATTTTATCTATT
TTAGAAGAATGTGGTTATGAAGCCAGTGCATTTGCGACCACACCAGAAGAAAATTCAGCACGCAATGAAGCACATCGTGCT
GCGCGGGCAGGATTTGATTTACTAGTAGCTGCAGGTGGAGATGGGACCATTAATGAAGTCGTGAATGGGATTGCTCCGTTG
AAGCGGCGCCCTAAAAATGGCTATTATTCCTGCTGGAACGACGAATGACTATGCACGGGCCTTGAAGATTCCTCGTGATAAT
ATCGTTAAGGCAGCAGAAGTGATTAAAAAAAAATCAAACTGTCAAAATGGATATTGGCCAAGCCGGCAAAAATTACTTTATC
AATATTGCGGCGGGTGGTCATTTAACGGAACTGACTTATGAAGTTCCGTCAGAGTTGAAAAGTATTTTTGGTTACTTAGCG
TACTTAGCCAAAGGAGCCGAAATGTTGCCGCGAGTGAAGCCGATTAAAAATGCGCATGACGTATGATGAAGGTGTGTACGAA
GGCAATGCATCAATGTTTTTCCTAGGGCTGACTAACTCGGTGGGTGGTTTTGAGCAAATCGTACCAGATGCTAAATTAGAC
GATGGTAAGTTTTCATTAATCATTGTAAAAACAGCCAATATTTTTGAGATTCTTCATTTAGTTGCGTTAATGTTAAATGGT
GGAAAGCATGTTGAAGATCATCGACTGATCTATACAAAGACCAGCTATTTACATGCAGAAACGTTAGAAAAGAACAATAAA
ATGATGATTAATTTAGATGGTGAATATGGAGGCGATGCTCCAATGACCTTTAAAAATATGCATCAACACATTGAAATTTTT
GCAAATGGTGATGCACTGCCGTCCAATGCAATTATGGGTTCTGTCTTAACTGGTAGCGATGAGATTGTCGTAGAATCAGAA
GACGAAGAGGAAGAAGCATATAACGAAGCCAGCAAAAGAATTTGTCAAAGAAGTTGAACGACTAACAGACGAAGATATTGAT
GGCGATGGAAAGATTGCGGAAAAAGAAAGCAC

EF0775
Seq ID 384
ATGACCAATACGGTGAAAGTGAAAGACGACAGTCTGGCTGATTGTAAACGGATATTGGAAGGACAAGCTACTTTCCCAGTT
CAAGCGGGTGAAACGGAACCAGTCGATTTAGTAGTTGTTGAAGATGCTAGTGGTAGTTTTTCAGATAATTTTCCACATGTA
AGACAAGCGATTGATGAAGTGGTTCAAGGCTTATCTGATCAAGACCGCGTGATGCTGGCTTCATATCGCGGCGGAAAACAA
TTTATGTTTCCTGATGGAAAGACAAAAATTAATTCAGCTGATTATGATATGAATGTGCGCGTCAATACGCAATTGACTTAT
GATAAAAGCCAATTTGTCTCTGGTTTTGGAGACGTTCGGACGTATGGTGGTACGCCAACCGCCCCAGGATTGAAACTCGCT
TTAGATACGTACAATCAAACACACGGAGATTTAACGAATCGAAAAACGTATTTCCTATTAGTGACAGATGGGGTCGCTAAT
ACACGTTTAGATGGTTACTTGCATAAGACCAATACCAATGATTCAATCAATGAATATCCAGATCCAAGCACATCCTCTTCAA
GTCTCAGTGGAATATAGTAATGACTACCAAGGTGCAGCAGCAGAAGTTTTAGCGTTAAACCAAGAAATTACTAACCAAGGC
TATGAAATGATTAATGCGTATTGGGAAAGTGTTGAATCTTTAAGTTCAGTGAATTCATACTTTGATAAATATAAAACAGAA
GTGGGTCCTTTTGTAAAACAAGAGTTGCAACAAGGGTCTAGCACACCAGAAGATTTTATTACAAGCCAATCTATTGATGAT
TTTACAACCCAATTAAAACAAATTGTCAAAGATCGTCTGGCGCAATCGACACCAGCAACAGCTTCATTAACGATTGCCAAT
CAATTTGATATTCAATCTGCGACGCCTACGGACGATGCTGGAAATGATGTGCCTGTTCAAATTAACGGACAAACCATTTCA
GCAACTAGTACAGAAGGTTACGTAGGAAACATCACGATTCACTACGAAGTCAAAGAAAAATACAGCGATTGATGCAGCAACC
CTTGTAAGTAGTGGGACAATGAATCAAGGAACAATTGCTAAGGAATTTCCAGAAGCGACGATTCCTAAAAAATGACAATGCG
CATGCGTGTGACGTGACGCCAGAAGATCCAACGATTACAAAAGATATCGAAAATCAAGAACACTTAGATTTAACCAATCGT
GAAGATAGTTTCGATTGGCATGTCAAAAACAGCCTTTGGCAACGAAACCAGTACTTGGACCCAAGCCAGCATGGTGGATGAC
ATTAATAAAGTGCTAGATATCATTGATGTGAAAGTCACCGACGAAAATGGTAAAGATGTTACAGCTAACGGCACAGTAACA
CAAGAAAATAACAAAGTAACTTTTGAAATGAACAAACAAGCAGACAGCTATGACTATTTAAGTGGTCATACGTATACAATG
ACTATCACCACTAAAATTAAAACTGACGCAACGGACGAAGAATTAGCGCCTTACATTGAACAAGGCGGGATTCCCAACCAA
GCCGACTTAAACTTTGGCAATGAAGGTGACGTGTTACATTCCAACAAACCAACCGTAACACCACCGCCAGTTGATCCAAAT
ATTGCTAAAGACGTAGAAGGACAAGAACATTTAGATTTAACCAACCGCGATCAAGAATTTAAATGGAACGTCAAAACAGCT
TTCGGTAACGAAACAAGCACTTGGACCCAAGCCAGCATGGTAGTGACATTAATAAAGTGTTAGACATCACTGATGTAAAA
GTCACAGATGAAAATGGTAAAGATGTTACAGCTAACGGCAAAGTAACACAAGAAAATAACAAAGTAACTTTTGAAATGAAC
AAACAAGCAGACAGCTATGACTATTTAAGTGGTCATACGTACACAATGACCATTACTACTAAAATCAAAGCTAGCGCAACG
GACGAAGAATTAGCACCTTATATTGAACAAGGTGGCATTCCCAACCAAGCCGACTTGGCAACGAAGGTGACGTG
TTGCATTCCAACAAACCAACCGTAACACCCACCTGCACCAACGCCAGAAGATCCAACGATTACAAAAGATATCGAAGGCCAA
GAACATTTAGATTTAACCAACCGTGACCAAGAATTTAAATGGAACGTCAAAACAGCTTTCGGTAACGAAACAAGCACATGG
ACCCAAGCCAGCATGGTGGATGACATTAATAAAGTGTTAGACATCACAGACGTGAAAGTTACTGATGAAAATGGCAAAGAT

```
GTTACAGATAATGGCATAGTAACACAAGAAAATAACAAAGTAACTTTTACTATGAACAAAAAAGATGACAGCTACTCTTAC
TTAGCTGGTCATACATACACAATGACTATTACCACTAAAATTAAAACTGACGCAACGGATGAAGAATTAGCGCCTTATATT
GAACAAGGCGGGATTCCCAACCAAGCCGACTTAAACTTTGGCAACGAAGGTGACGTGTTGCATTCCAACAAGCCAACCGTA
ACACCGCCTGCACCAACGCCAGAAGACCCAAAAAAACCTGAACCTAAACAACCGCTAAAACCGAAAAAACCGTTGACGCCT
ACAAATCATCAAGCACCAACGAACCCAGTCAATTTTGGAAAATCAGCAAGTAAAGGAATTCATTTACCAATGACTAATACA
ACAGTAAATCCACTTTACATGATCGCAGGTTTAATTGTCCTTATAGTGGCTATTAGCTTTGGCATAACAAAAAATAAAAAA
AGAAAAAAT
```

EF0779

Seq ID 385

```
ATGTCTGCTTGGCAAACATTTAAAAATGGAACAAGGCAATTTTTTAAAGACATCTTGCAGTATTTATGGTTATTTTTTACT
TTAAATGTCTTATTATTATTGGTGGGTGGGGCTTTTAGTTGGGCTACTTCCAACGCACTAAAAACTCAAGGTATTCCTTAT
CTTTCGTTTAATAACCTGAATTTATTGTTAGAAAAGCCGCTTGCTCTGGTGTTATTAATTCTCTTGTTGTTACTTTTTTTA
GGTGCAGTTTTTTATCAGTTTACCTTTTTATTATTAGGGATTTTTCAAATACGGCAAGACCATCGCTTTCATTTTAAAGGA
GTTACCAAGGCATCTTTTAAGGTTCTTAAAAAGCAAGGTGCCCGTTCGTGGTTATTCTTTTTCGGTTATTTTGTTGTCATT
GTCCCTTTTGGTAACCTGATTTTCCAAAGTAATCTATTGACGAAATTTGTGATTCCAGATTTTATTGTTGAATTTTTGTCC
CAACGAATTCCTTACTTGGTGGGGTTACTGGCCTTTGGGTTATTAGTTCTGGTATTTAGCTATTCGTTTCATTTACACGCTT
CCATTAATGATTCTGGAGAGAAAAAAAGCGGGCGAGGCAGTTAAGGCTAGTTGGTCAATGACAAACAAACGATTATGGTTC
ATTATTCGTAATATTGCATTTGTGACGATCGCGGTATTTGTTTCAACGTATGTTATTTATGTGTTACTTTATCTTCTACAG
TTGAAGTTAGATACCTTGAGCGATACCATTAGCTTGCTAGGTGGGATTTTAAATTTAACCGTTGTCCAATTTTTTACAGTTT
GTCTCTAATGCTTGGCTTTCAGTATTACTGATTAATTTCTTATATACACAATTAAATGTTCAAGCAGAAACGACGACAAAA
GTTGCCTTTGACAAAGAAACGAAACGTAATAAAATTAGTGACTATCGGGATGGGCCTAGGACTATTCACCATTTTTGGCGGC
TATATAATTTTCAACGCTGTTTACTTAACAGGGCTATTAGAATCAAAACCTTTGATTATTTCGCATCGAGGGGTCACCAAT
AGTAATGGCGTACAAAATACCATTCCTGCTATGGAAAGAACCATTAAATTTAAACCAGATTATATTGAAATTGATGTTCAA
GAAACCAAAGACCATCAATTTGTTGTGATGCATGATGCAAACCTTCAAGAACTCGCCGGCCGTTGATGGCACGCCGCAAGAA
TTTACGTTGGCTGAATTAACAAAAATGACCGTTAAAGAAAATGGCCAAGAAGCACCCAATTGCCAGTTTTGATGATTATTTA
GCGAAAGCAAACCAAGCCAAACAAAAATTACTAGTAGAAATTAAAACGTCTAAACAAGATAGTCAAGGTGCGCTTTCTAAC
TTTATTGAAAAATATGAACGTCCTCTGATAAAGAATAACCATCAGGTTCAATCTCTCGACTACAATGTGATCAAAGCGTTT
AAAAAAGCCAAATCAAAAGTAAAAGTTAGTTTTATTTTGCCCTATAATTTCACTTTTCCAGAAACACAAGCTGATTGTAT
ACAATGGAAGCTACTACGTTGAATGATACCTTCATTTTAAAAGCTGACCAACAGAAAAAAGCAGTTTATGCTTGGACAGTT
AATGACTCAGAGGTATTGAGTAAAAATGCTCTTTATGGACGTTGCTGGAGTGATTACCGATGATTTGGAGTTAGTCAATGAA
GAAGTCAATGACTTTGAAAAAAATCCGTCGTATGCGGATCGTATTTTGCATTATATTTTTATGCTACCAAGTGTGGCTTCT
CAA
```

EF1091

Seq ID 386

```
ATGATAACAGATGAGAATGATAAAACGAATATTAATATCGAGTTAAATCTTCTCAACCAAACAGAGCAGCCATTACAACGA
GAAATTCAATTGAAAAATGCACAGTTCATGGATACTGCTGTAATTGAAAAAGACGGATATTCTTACCAAGTGACTAATGGT
ACGCTTTATCTGACTTTGGACGCACAAGTAAAAAAGCCGGTACAGCTTTCGTTAGCTGTTGAGCAAAGTTCGCTTCAAACA
GCTCAGCCACCTAAGTTATTGTATGAAACAACGAATATGATGTTTCAGTTACTTCTGAAAAAATAACAGTAGAGGATTCT
GCTAAAGAATCAACTGAACCAGAAAAAATAACTGTACCAGAAAATACGAAAGAAACTAACAAAAATGATTCGGCTCCAGAA
AAAACAGAACAGCCGACCGCAACAGAGAGGTAACCAATCCATTTGCAGAAGCAAGAATGGCGCCAGCTACTTTGAGAGCG
AATCTGGCACTGCCTTTAATTGCACCACAATACACGACGGATAATTCTGGGACTTATCCGACAGCTAATTGGCAGCCCACA
GGCAATCAAAATGTGTTAAACCATCAAGGGAATAAAGACGGTAGTGCACAATGGGACGGCCAAACGAGTTGGAATGGGGAC
CCTACTAATCGCACAAATTCTTATATTGAGTATGGCGGTACAGGAGACCAAGCCGATTATGCCATCCGAAAATATGCTAGA
GAAACAACAACACCAGGGCTTTTTGATGTATATCTTAATGTGCGTGGGAATGTTCAGAAAGAAATCACGCCATTGGATTTG
GTCTTAGTCGTTGACTGGTCCGGTAGTATGAATGAAACAATCGGATTGGTGAAGTTCAAAAAGGAGTGAACCGTTTTGTT
GATACATTGGCAGATAGCGGTATTACCAATAACATCAACATGGGCTATGTTGGCTACTCAAGTGACGGTTATAATAACAAC
GCCATTCAAATGGGGCCGTTTGATACAGTCAAAAATCCAATTAAAAATATTACGCCAAGTAGCACTAGAGGAGGAACTTTC
ACTCAAAAAGCATTAAGAGATGCTGGTGATATGTTAGCAACGCCAAATGGACATAAGAAAGTCATTGTACTTTTAACGGAT
GGCGTCCCAACCTTCTCTTATAAAGTGAGTCGAGTTCAAACAGAGGCGGATGGTCGCTTTTACGGGACACAATTTACGAAT
CGACAAGATCAACCAGGTAGCACTTCTTATATCTCTGGTAGCTATAATGCGCCAGATCAAAACAATATCAATAAACGGATT
AACAGTACGTTTATCGCCACGATAGGTGAGGCAATGGTCTTAAAACAACGTGGGATTGAAATACATGGATTGGGCATTCAA
TTGCAAAGCGATCCACGAGCTAATTTATCTAAACAACAAGTTGAAGATAAAATGCGTGAGATGGTGTCAGCCGATGAAAAT
GGAGACCTTTATTATGATCCGCGGATTATGCACCGGACATTTCACCGAGTATTTTAGCGAAAAAAGCCGTTCAGATTTCAGGA
ACGGTTGTAAACGGAAAAGTAGTTGATCCAATTGCTGAACCTTTTAAAATACGAGCCAAATACATTATCAATGAAAAGTGTG
GGTCCTGTTCAGGTTCAAACATTACCAGAAGTGTCGCTAACAGGCGCTACAATTAATAGTAATGAGATTTATTTGGGTAAA
GGGCAAGAAATTCAAATTCATTATCAAGTACGTATTCAAACAGAGTCAGAAAACTTCAAACCTGATTTTTGGTATCAAATG
AATGGTCGGACAACGTTTCAGCCATTAGCCACGGCCCCTGAAAAAGTTGATTTTGGGGTTCCTTCGGGAAAAGCACCTGGC
GTGAAGTTAAACGTGAAAAAAATCTGGGAAGAGTATGATCAAGACCCGACAAGTCGGCCAGATAATGTGATTTATGAAATT
AGTAGAAAGCAAGTAACTGACACAGCCAACTGGCAACTGGGTATATTAATTCATCAAAATCAAGAAAATGATACCAGCAAT
AGTTGGGAGCGCAAAAATGTAACCCAACTTTCCAAAACCGCGGATGAAAGCTATCAAGAAGTTCTTGGGCTTCCCCAATAC
AACAATCAAGGACAAGCTTTCAATTATCAAACAACCCGTGAATTAGCAGTTCCTGGTTACAGTCAAGAAAAAATCGACGAT
ACTACTTGGAAAAACACGAAGCAGTTCAAGCCATTAGATTTAAAAGTAATCAAAAATTCTTCCTCAGGTGAGAAAAACTTA
GTGGGAGCCGTCTTTGAATTGAGTGGTAAAAATGTTCAAACAACATTAGTGGACAATAAAGATGGTAGCTATTCCTTGCCA
AAAGATGTGCGCCTACAAAAAGGGGAACGCTATACATTAACTGAAGTAAAAGCACCTGCAGGACATGAGTTAGGCAAGAAA
ACGACTTGGCAAATTGAGGTGAGTGAGCAAGGCAAAGTAAGCATCGATGGACAAGAAGTGACCCACACAAATCAAGTTATT
CCATTGGAAATTGAAAATAAATTTTCTTCTTTGCCAATCAGAATTAGAAAAATACACCATGCAAAATGGCAAACAAGTGAAC
TTAGCAGAGGCGACTTTTGCGTTGCAAAGAAAAAATGCTCAAGGAAGTTACCAAACTGTGGCAACTCAAAAAACAGATACT
ACAGGATTGAGCTATTTTAAAATTAGTGAACCTGGTGAGTATCGAATGGTGGAACAATCAGGACCATTAGGCTACGACACT
CTTGCTGGAAATTATGAATTTACTGTTGATAAATATGGGAAAATTCACTATGCAGGCAAAAATATTGAAGAAAATGCGCCA
```

GAATGGACACTGACACATCAAAATAATTTGAAACCTTTTGACTTAACAGTTAATAAAAAAGCCGATAATCAGACGCCACTT
AAAGGAGCGAAATTCCGTTTAACAGGACCAGATACGGATATTGAATTACCAAAGATGGCAAAGAAACGGATACTTTTGTT
TTTGAAAACTTAAAACCAGGGAAATATGTTCTAACAGAAACCTTTACGCCAGAAGGATATCAGGGGTTAAAAGAACCAATC
GAATTAATAATTCGTGAAGATGGTTCAGTCACGATAGATGGGGAAAAAGTAGCAGATGTTTTAATTTCTGGAGAGAAGAAT
AATCAAATTACTTTAGACGTTACGAACCAAGCAAAGGTTCCTTTACCTGAAACTGGTGGCATAGGACGCTTGTGGTTTTAC
TTGATAGCGATTAGTACATTCGTGATAGCGGGTGTTTATCTCTTTATTAGACGACCAGAAGGGAGTGTG

EF1323
Seq ID 387
ATGGGGATTTTGTTACCGATTATTATTGCAGTTTTAGTATTATTAATGTTATTAATTGTGTTTGTTTCTAAATATCAGACA
GCCAAACCTGATGAAGCGTTAATTATCAGCGGGAGCTATCTAGGCTCTAAAAATGTTCATGTAGACGAAGGTGGCAACAAA
ATTAAAATCGTTCGTGGCGGTGGTGCGTTTGTCTTACCAGTGTTCCAACGTTCAAATCGAATTAGTTTGCTTTCAAGTAAA
TTAGATGTTTCCACACCAGAAGTGTACACAGAACAAGGGGTTCCTGTGATGTGTGATGGGACATCAATCATTAAAATTGGT
TCGTCAGTTGAAGAAATTGCGACAGCAGCGGAACAATTTTTAGGAAAAACAACGGAAGAATTAGAAAATGAAGCACGTGAA
GTATTAGAAGGACATTTACGTTCGATTTTAGGTTCAATGACAGTGGAAGAAATTTACCAAAATCGTGATTAATTTAGCCAA
AGTGTACAAGAAGTTGCCAGTGTTGACTTAGCTAAAATGGGCTTAGTTATTGTGTCGTTCACAATTAAAGAAGTTCGTGAT
AAAAATGGATACTTGGATTCATTAGGGAAACCAAGAATCGCTCAAGTTAAACGTGATGCAGATATTGCAGAAGCAGAAGCC
TTGAAAGAAACTCGCATCAAAAAAGCAGAAGCGAAAAAGAATCACAACAAGCGGAATTGCAACGTCAAACAGAAATTGCA
GAAGCTTCAAAAGAAAAAGAATTGAAATTAGCGTTATACAAACAAGAACAAGATATTGCCAAAGCCAAAGCAGACCAAGCC
TATAACTTGGAAAGCGCGCGTGCACAACAACACGTGGTGGAACAAGAAATGGAAGTCAAAGTCGTTGAACGTCAAAAACAA
ATTGAGTTAGAAGAAAAAGAAATTACGCGTCGTGAAAAAACAATACGACTCAGAAGTGAAGAAAAAAGCCGATGCAGATCGT
TACGCACGCGAACAAGAAGCTCTTGCTCAAAAGGCACGCGAAGTGGCAGAAGCCGAAGCAGAACGCTTCAAAGTTGAAGCA
TTAGCCGAAGCGGAAGCCAACAAAACGCGTTTAACCAGGTCAAGGACGCCGAAGCTATTTTAGCTCGTGGTGCAGCGGAA
GCCGAAGCCAAACAAAAGATTGCTGATGCCTTCAAAGAATATGGTGAAGCAGCTGTGTTAAGTATGGTGATGGAAATGTTG
CCACAATTAATGAAAGAAGCAGCACAACCATTGGGCAACATCGACAAGATTTCAGTGGTAGATACAGGCGCAGGTGGCGAA
AATTCTGGCGCCAACCGTATTACAAATTATGCAACAAATTTGTTGGCTGGTACACAAGAAACACTAAAAGAAACAACGGGC
CTCCGACGTGAAAGAGCTAATTGAAAACTTTTCTAAAAAAGGTACCTCAAATAGCGTGAACTATCATGCAACAGAAGGTTCA
GAAAAAGAA

EF1355
Seq ID 388
ATGGCTTATCAGTTTAAATTACCGGATATCGGTGAAGGGATTGCCGAAGGCGAAATCGTTAAATGGTTTGTAAAACCTGGC
GATACAATCAACGAAGACGATACGTTATTAGAAGTACAAAATGACAAATCAGTGGAAGAAATTCCATCACCAGTAACAGGT
ACTGTAAAAAATATCGTTGTACCAGAAGGAACAGTTGCAAACGTTGGTGACGTGTTAATCGAAATCGACGCACCTGGTCAC
GAAGATAACGATGCAGCACCAGCAGCTCCTGCACAAGAACAAACACCAGCACAACCTGCTGCTGTACCAACAACCGAAGCA
GCTGGCGGATTTTTCCAATTCAAATTACCAGACATCGTGAAGGAATTGCCGAAGGCGAAATCGTTAAATGGTTCGTTAAA
GCGGGCGACACAATTAATGAAGATGATTCATTATTAGAAGTACAAAATGACAAATCAGTAGAAGAAATTCCATCACCAGTA
ACAGGTACTGTAAAAAATATCGTTGTACCAGAAGGAACAGTTGCCAATGTGGGTGACGTGTTAGTTGAAATTGACGCACCT
GGTCATAATTCAGCAGCACCGTCAGTCGCAGCACCAGCTACTGACGCTCCTAAAGCGGAAGCATCAGCTCCAGCCGCTTCA
ACAGGCGTAGTTGCAGCCGCTGATCCAAACAAACGCGTTTTAGCAATGCCATCTGTTCGTCAGTATGCGCGTGAAAAAGAC
GTTGATATTACACAAGTAACTGCAACTGGTAAAGGTGGCCGTGTCATTAAAGCGGATATTGATGCCTTTGTCTCTGGTGGT
TCTCAAGCAGCCCCAGCTACTGAAGCTGCCGCAACAGAAGCGCACCTAAAGCGGAAGCAGCTGCACCTAAAGCAGCGCCA
AAAGCCTTTACTTCTGATTTAGGCGAAATGGAAACACGTGAAAAAATGACACCAACACGTAAAGCAATTGCTAAAGCAATG
GTTAACAGCAAACACACTGCTCCTCACGTAACATTACATGATGAAGTAGAAGTTTCTAAATTATGGGATCACCGTAAGAAA
TTTAAAGATGTTGCTGCTGCAAATGGTACAAAATTAACATTCTTACCATACGTTGTAAAAGCATTGACTTCAACTGTTCAA
AAAATTCCCAATCTTGAATGCATCAATCGATGACGCAGCACAAGAAATTGTTTACAAAAATTACTTTAACATTGGTATCGCT
ACTGATACAGATCATGGCTTATATGTACCAAATGTTAAAAAATGCTAATACGAAGAGCATGTTTGCTATCGCTGATGAAATC
AACGAAAAAGCAGCATTGGCTATCGAAGGTAAATTAACTGCACAAGATATGCCGTGATGGTACAATCACAATTAGTAACATT
GGTTCAGTCGGTGGCGGCTGGTTTACACCAGTAATCAACTACCCTGAAGTTGCTATTTTAGGCGTTGGTACAATTGCACAA
GAACCAGTTGTTAATGCAGACGGCGAAATCGTTGTGGGACGCATGATGAAATTATCATTAAGCTTTGACCACCGTATCGTT
GACGGCGCAACTGCTCAAAAAGCAATGAACAACATTAAACGCTTATTAGCTGATCCAGAATTACTATTAATGGAAGGA

EF1699
Seq ID 389
ATGGAATATACAATTAAAAAAATGGCTTCACTATCTGGCGTCAGTGCACGGACGTTACGATATTATGATGAAATTGGTCTT
CTTCAACCAGCCAGAATTAATTCTTCTGGTTATCGAATTTACGGGCAAGCCGAGGTGAATCGTTTGCAACAAATTCTTTTT
TACCGTGAATTGGATCTGAAACTAGATGAAATTAAAGAAATTTTGGAGCAACCTGACTTTAATGTTGAACAGGCTTTATAC
GAGCATCAACAAAAACTATTGGAAAAGCGCAATGAAATTGATCGTCTTTTAGCCTCTGTTCAACAAACCTTACACCATTAC
AAAGGAGAGATAAACATGTCAGATCAGCAAAAAATTCGAAGCATTCAAACAACAAAAAGTTCAAGAAAACGAAGAAAAATAT
GGTAAAGAGATTCGGGAAAAATATGGCAACGAAACAATTAACAAGCCAACAAAAAAATATTTAAACTTGACGGAAAAAGAC
ATGCAAGCAATGCAGAACGTGGAAAAAGATTTATTTTCTAAGTTAGCCATGTATCAAAAATCACCCAAGTTGACTAGCCAA
CTTGCGCAAGAAATTTTCCAATTACACAAAGATTGGTTAATGTACTCTTGGTCAAGCTATTCGCCAGAAGCACACAAAGGG
TTAGGCCTTATGTATGTTGGCGATGAACGATTTACGTCATATTATGAACAACACGGCGCTGGTTTTGCCGAAGCGTTAAAC
GCAATTATCCAAAACTACGCT

EF1744
Seq ID 390
ATGGCTAAAAAAGGCGGATTTTTCTTAGGAGCAGTAATTGGTGGAACAGCAGCAGCCGTTGCCGCATTATTACTTGCACCA
AAATCAGGTAAAGAATTACGTGATGATTTATCAAATCAAACAGATGATTTAAAAAAACAAAGCGCAAGATTACACAGATTAT
GCTGTTCAAAAAGGAACAGAATTAACAGAAATCGCAAAACAAAAAGCCGGCGTTTTATCAGATCAAGCCTCTGATTTGGCA

GGTTCTGTCAAAGAAAAAACAAAAGATTCATTGGATAAAGCACAAGGTGTTTCTGGCGACATGCTTGATAACTTTAAAAAA
CAAACAGGTGATTTATCTGATCAATTTAAAAAAGCAGCTGACGATGCTCAAGATCACGCGCAGAAGATTTAGGTGAAATTGCC
GAAGATGCAGCAGAAGATATCTATATTGACGTTAAAGATTCTGCGGCAGCGGCCAAAGAAACTGTTTCTGCTGGTGTCGAT
GAAGCAAAAGAAACCACCAAAGATGTTCCTGAAAAAGCTGCAGAAGCAAAAGAAGATGTTAAAGATGCAGCGAAAGACGTA
AAAAAAGAATTTAAAGGG

EF1752
Seq ID 391

ATGAAAAGAAAACTAACAAAATCTCCGAATAACGTGGTTTTAACAGGAAGCTTAGCTGGGATTGCTGATTGGCTAGGAATT
GACCCAACAATTATTCGTGTCGTTTATGTCTTGCTCAGCTTTTTTCTCAGCAGGATTTCCAGGCATTCTCTTATACATTGCT
TTAGCAGTCTTGATCCCTTCTGGAAGAACAGGAAGCGATCGCGGCTACGGCCATCAAAATCCTTATAATAGAAATGTTCGC
AATGAAAATCCTTATGCTGCCAATAAAAAGCAACGTAAAGAAGCGGAAAAAAATTGATGATGACGAATGGAGTGACTTT

EF1753
Seq ID 392

ATGAAAGAAAGAGAACGCGTATTAGAATTAGTGAAAAAAGGTATTCTAACGTCAGAAGAAGCGTTAATTTTATTAGAAAAT
ATGGCAACTGAAAAAGATGAAAAACAAATCGAAAAAGCTGCTGAAAAAGTTGATACACAAAATATTGGAACAACAAATAAA
GAAGATCAAGTCGCAGATTTAATGAATGCATTAGAAAAAGGCGAATCAGAAGGACCTACTGTTGATTCGTTTGAAGAAAAT
ACACAGGACAGTGCAGAAAAAGATCGTGAAAACTTAGAAAGAATTCTTGATGAGTTAGCAACAAAAGCCAATCGTGCTTCG
GCTGAATTGGACGAAGTCAACGCAGAAATTGCCGGCATCAAAGAAGAAATTAAAGAAGTCGCAGAAGAAATTGGAACATTA
GATACAAAAGAAGAATTAGATGCATTAACAGAAGATGAACAAGTTCAACGAAAAGACTTGCACGTTTTACTTGCACAATTA
GAAGAAAAATTAGCGACTCAAAGTACTGAAAAAACAGCACTGGAAGAAGAACTAAAAAACATTCGCAAAGAACAATGGAAA
GGTCAATGGAATGATACAAAAGAAAAAGTTTCTTCTCAATTCTCTGAAGAGTGGAAAGATCAAGCCACAGACACCTTTAAC
CAAGTCGGCGGCAAAGTTGCCGAAGTTGGTGGTCAAGTGGGAGAATTCTTGAAAAAAACATTTAATTCTTTCAGTGATACC
ATGAATGATAATGTGGAATGGAAAGACATTAAAATGAAAGTTCCTGGTGTGGCAACAACTAAGTTTGAACATGAGTTTAAC
TATCCAAATCCACAAGCAAGTTTAATTGATGTCAAGGTAGCAAATGGTACCGTGGTTTTCAAAACTTGGGATCAAGAAGAT
GTGAAAGTCGAAGCAAAAATCAAATTATATGGTAAAATGGCAGGAGATTCACCAATGGAAGCTTTCTTAGAACGAAGTGAC
ATTGATGTGGATGATGAAACGATTTCTTTCCAAGTGCCAAACAAACGGGTGAAAGCAGATTTAACGTTCTATTTACCAAAA
CGCACTTACGATCATGTATCTGTTAAATTATTAAACGGAAATGTCTTAGTAGAAGAGTTAACAGCGAAAGATGTTTACACA
AAATCAACGAATGGGACGATTACGTTTAAAAAAATTGAACCTCAACTATGTTAGAAATTGAAGGTGTGAATGGTGAAATTAAA
GTCCTAGAAGGAACGATTTTAGATAACATTATTGAAACAGTCAATGGTGATGTGTCTATTTCCGCGGCACCAGAAAGTCTA
AGTGTCTCTCTGATTAATGGAGATATCCGCATTACAGCTAAAGAAAAAACGCTTCGCCGTGTTGAAGCAAGCTCTGCAAAT
GGGAATATCAAATTGGCCTTGCCAAATGACTTAGGTGTTGAAGGGCAAGTGAAAACTAATTTAGGTAGTATTAACAGCCGT
TTAACAGATATTGAAGTTGTTCGTGAAAAGAAAGATCGCGGCAATCAACAATTACATTTTAGACGTGTACTGGAAGAATCA
ATGGGCTCAAATTAATGCTTCTACAACAACGGGAAGTATTTTCCTAAAAGATACGGATAAA

EF1791
Seq ID 393

ATGAAAAAGTATTTAAAAAATCACAATGGTTTGTATTTTATTGGTAGGATTTTTAGCTGGGGTGTACCAATAAAAATGAAAAT
AAAAAGAAACAGAAAAATACCAAAGAAGCCGTTCAACTGATGTCACCCTCGGAATTAACAACGCTCAACACCTCTGTATTA
TTGGATTTTCCAGATGCTATTGTCCAAACTGCAGCGTTTGAAGGGTTATATAGTTTAGATGAACAAGACCAATTGGTACCA
GCCGTAGCAAAAGCATTGCCGATGATTTCAGAAGATGGAAAAACCTACACGATTTCTTTGAGAAAAGAAGCGGTTTGGAGT
AACGATGATCCTGTCACAGCACATGATTTTGAATATGCTTGGAAAAAAAATGATTGATCCTAAAAACGGCTTTGTTTATAGC
TTCCTCATCGTTGAAACAATTCAAAATGGTGCAGAAATCTCAGCGGGGAAATTAGCACCCAATGAACTAGGTGTCACAGCT
GTGGATGATTATACATTAAAGGTGACGCTCAAAGAGCCAAAACCGTACTTTACGTCCTTGTTAGCTTTTCCGACATTTTTC
CCGCAAAATCAAAAAGTAGTCGAACAATTTGGTGCGGACTATGGAACTGCTAGTGATAAAGTCGTCTATAATGGTCCGTTC
GTGGTAAAAGATTGGCAGCAAACAAAGATGGACTGGCAACTAGCAAAAAATAATCGCTATTGGGATCACCAGAACGTGCGC
TCAGACATTATCAATTATACAGTTATCAAAGAAACATCTACCGCATTGAATCTTTTTGAAGATGGACAATTAGATGTGGCT
ACACTAAGTGGTGAACTGGCGCAACAGAATAAAATAATACGTTGTATCATTCGTATCCAACAGCGACAATGAACTATTTG
CGCTTAAATCAAAAACGGAAAGGGCAAGCAACGCCGCTTGCAAACGAAAACCTGCGTAAAGCATTGGCTTTAGGAATAGAT
AAAGAAAATCTAGTCAATAATATTATTGCAGATGGTTCTAAAGCGCTACATGGTGCGATTACGGAAGGCTTTGTGGCGAAT
CCCACAACGGGGTCTCGATTTTCGTCAAGAAGCAGGTAATTTAATGGTTTATAACAAAGAAAAAGCGCAAAGTTATTGGAAA
AAAGCACAAGCAGAATTAGGAGAAAAGGTTAACGTTGAATTGATGGTAACAGATGATGGTTCTTACAAAAAAATTGGTGAA
AGTTTGCAAGGCTCGCTACAAGAATTGTTTCCTGGTTTGACAATAGAGCTAACCGCATTGCCGACTGAAGCTGCATTGAAC
TTTGGGCGAGAAAGTGACTATGATTTATTCTTAATTTACTGGACACCAGACTATCAAGACCCTATTTCTACCCTGATGACT
TTATACAAGGGCAATGATCGCAATTATCAGAACCCTGTCTATGACAAATTATTAGATGAAGCAGCCACAACCTATGCCTTA
GAGCCAGAAAAAAGATGGGCGACACTGATTGCAGCTGAAAAAGAAGTGATTGAAACGACTGCTGGCATGATTCCACTTAGC
CAAAATGAACAAACAGTCCTGCAAAATGATAAAGTCAAAGGCTTGAATTTTCATACCTTTGGCGCTCCATTAACGTTAAAA
AATGTTTATAAGGAAAAA

EF1800
Seq ID 394

ATGAAACATGGAAAAATAAAACGATTTAGTACATTGACACTATTGGCAAGCGCAACGATTTTAGTACCATTAAGTACGTCT
GCAGAAGAAACAACAAACAGCAGCACTGAAACAAGCTCATCAATGGTGGAACCTACCGCAACAGAAGAAAAATTGTGGCAA
TCTGATTTTCCAGGAGGGAAAACTGGTGAGTGGCAAGATGTGATTGGTAAAACAAATCGAGAATTAGCAGGAGAGTCATTG
GCGATTTCACGAGATGCAGCAGCAGGTAATAATGCCGTATCTTTAAATTTAGATTCCCCTAAATTAGCTGATGGTGAAGTA
GAAACAAAGTTTAAATACACTGCTGGAAGCGGTCGAACGGGTGTGATTATACGTGGTAATAACCAAAGATAGCTGGGTTTTT
GTAGGCTATAACGCGAATGGCAAATGGTTAGTTGAAAGCCGTAATTCGTGGAATGATTCAATTTCTGGGCCAACGTTAAAT
GAAGATACGAATTATTTGTTGAAAGTACGTTATGTTGGTGAAAAAATTACTATCTGGCTTAACACCACGTTGATTTACGAA
GGAGAACCTGTTTTGGCCAATGGAGACAAAATTCCAACAGAAGCCGGACATGTGGGTGTCCGTTTGTGGTACGACAAAAAA
ATTGTCAATTATGACTATTTTAAAAATGGCCCCGTAGATAGCATTCCAGAAATTGTGCCAGAAGTGACACAAATTGCGCCA

```
GTCAAAGTTTTTACAAAAATTGGTGTCGCACCAAAATTACCGAAACAAGTAAAAGTGACCTATAATACTGGTAAAGAAGCC
AATGAAGCAGTCCGTTGGAATGAAATCGATCCTGATGCATATAAAGAACCAGGAACTTTTTGAAGTCGACGGTACTTTGGAA
AATACAAACATCAAAGCAAAAGCCAGCATTGTTGTTGCTAAAGACAATGAAGCTGAAAAAGGCGACAAAATCTCTTCCGCA
GATTTAACAGCCGTGGTTGATCCACAATTTCCACGAATTATTCGCTACGAAGACCCTCAGAGTAATCAAGTGATTTTTAAT
GGCCAACACGAGAAAATTGACCAAGTAATGATTGATGGCAAGCATATAAAGCAACTGCTGAAAAACAGAAGAGTGAAGCA
AATCAAGCCGTTTATAACGTAGCTGTCCCAGAAATTGGTTTGCGTTTCACAACGACATTGACTGTTTCCGAAGGCCAAGAA
TTAGCTATGAAACTCTCAGATATTCGTGAAGAAGGAACCAAAATACACACAATTTCAATTCCAAATCAAGGCTTGATTTCT
GTCAATAGTACAGATGAAGGGGCGACTTTTGCTGGCGTTGTGATGAATACTGGGACAAATGCAAATAACGGAAATAAAAAT
GGTGATACTATCCAAGATTTAACTACAACAAGCCAAGAAGAAACGAAAAAATATATGTATGGTTTCTTAAATACGGCGAAT
TATGCTGCAAGTTTTTGGACGAACGCCTATGGAGACGGCTCTGTCGATGGTAGTGACAACAATCGAATCCATAAACAAACA
AAAGAAGCGGCGACTGGTTTTGTAACAACCTTGTCAAGTGGGGCATGGACCTATCGACCATTTGATCGCACCGGAAGATTAC
ACAACTGGAGAAACGCCAGAAGTGAAAGTTAAATTCTCAAAAGATAGCAACGACGACAATCGGGTGGATTGGCAAGATGCG
GCAATTGGGTTCCGTTCAATTATGAATAACCCAATGGGTGCGGAAAAAGTCCCTGAATTAGTCAACCAACGGATTCCTTTT
AACTTTGCTAGTCAGGCGACAAACCCATTCTTAGTGACGTTAGACGAATCAAAACGTATTTACAATTTAACAGATGGATTA
GGACAAATGAATTTACTAAAAGGGTATCAAAATGAAGGACATGATTCTGCGCATCCAGATTACGGTGCTATTGGTCAGCGA
CCTGGTGGGGAACAAGCGTTGAATCAATTAATTGATGAAGGACATAAATTAAATGCCGTTTTCGGTGTGCATATTAATGAC
ACCGAGTCTTACCCAGAAGCAAAAGGATTTAATGAGGAATTAGTTGATCCAACGAAGCGTGGCTGGGATTGGTTAGATCCG
TCTTATTTTTATTAAACAAAGACCCGATACATTGAGTGGTCGTCGCTATGAGCGCTTTAAAGAATTAAAACAAAAAGCACCG
AATCTAGATTATATTTATGTTGATGTTTGGGGCAACCAAGGCGAATCAGGTTGGGCAAGTCGTCAACTAAGTAAAGAAATT
AATTCACTCGGTTGGTTTACAACCAATGAATTTCCGAATGCTTTAGAGTATGACTCGGTTTGGAACCATTGGTCTGCAGAA
AAAGATTACGGCGGTACAACAACGAAAGGCTTTAACAGTACAATCGTTCGTTTTATTCGGAATCATCAAAAAGACACTTGG
ATTATTTCCGACAACCCTTTGTTAGGTGGAGCTGAATTTGAAGCCTATGAAGGTTGGGTGGGTAAAACCAATTTTAATACC
TATCGCCAAAAAACTTTTGCCATTAATGTCCCAACTAAGTTCTTACAACATTACCAAATTACAAACTGGGAAACTACAACA
GCAGCAGATGGTCAAATCTATGGCACAATTAAATTAGCGAACGGTGCTGAAAAAGTGACCGTTACTCAAGCAGATGCTAAT
TCGCCAAGAAGCATTACGTTAAATGAGACAGAAGTTCTAAAAGGTGATGCGTATCTACTGCCTTGGAATGTCAATGGTCAA
GACAAACTATATCACTGGAATCCAAAAGGCGGTACCAGCACTTGGTCATTGGATAAGAAAATGCAAGGAAAAACGAATTTA
CATTTATATGAATTAACAGATCAAGGGCGTATTGACAAAGGCGCAATTGCCACTACAAATAACCAAGTGACCATCCAAGCC
GAGGCTAATACACCGTATGTCATTGCTGAACCTGACAGTATTGAACCGATGACATTTGGAACAGGAACACCATTTAAAGAT
CCTGGATTTAATGAAGCCAATACCTTAAAAAATAACTGGAAAGTTTTCCGAGGTGATGGAGAGGTTAAAAAAAGATGCCAAT
GGTGATTATGTCTTTAGTTCAGAAAAAGAAAGAACCGAAATCAAGCAAGATATCAATCTTCCTAAACCAGGAAAATATAGT
TTGTATCTAAACACAGAAACACATGATCGTAAAGCCACAGTAACTGTTAAAATTGGTGGTAAGAAATATACGCGGACAGTG
AATAATTCGGTTGCCCAAAACTACATTCAGGCAGATATTAACCATACAAGCAGGAAAAATCCGCAGTATATGCAAAATATG
CGAATTGATTTTGAAATCCCAGATAATGCCAAAAAAGGCTCGGTGACATTAGCGGTTGATAAAGGCAATTCCGTTACAAAA
TTTGATGATTTACGAATTGTTGAGCGTCAAACGGATATCATGAACCCAGACAAACAAACAGTTATTAAGCAAGATTTTGAA
GACACACAAGCAGTTGGGTTATATCCGTTTGTTAAAGGCTCAGCTGGTGGTGTAGAAGATCCACGGATTCATTTATCAGAA
AGAAATGAACCTTACACACAATATGGTTGGAATGGAAACCTTGTTTCAGATGTATTAGAAGGCAACTGGTCCTTGAAAGCC
CATAAACAAGGAGCAGGATTGATGCTTCAAACAATTCCGCAAAATATTAAATTTGAACCGAACAAGAAATATACGGTCCAA
TTTGATTATCAAACTGATGGTGAAAATGTCTTTACTGCTGGGACCATTAATGGGGGAGTTGAAAAATAACAATGACTTTAAG
CCAGTCGGTGAGTTAACTTCGACAGCAGCAGATGGTCAAACCAAGCATTATGAAGCAGAAATAATTGGGGATGCTTCAGGA
AACACTACGTTTGGTATTTTTACAACAGGTGCCGATAAAGATTTCATTATGGATAACTTTACGGTCACAGTGGAATCAAAA
AAA
```

EF1818
Seq ID 395
```
TTGATGAAGGGAAATAAAATTTTATACATTTTAGGTACAGGCATCTTTGTTGGAAGTTCATGTCTATTTTCTTCACTTTTT
GTAGCCGCAGAAGAACAAGTTTATTCAGAAAGTGAAGTTTCAACAGTTTATCGAAGTTGGAAAAGGAGGCAATTTCTGAG
GCAGCTGCTGAACAATATACGGTTGTAGATCGAAAAGAAGATGCGTGGGGGGATGAAGCATCTTAAGTTAGAAAAGCAAACG
GAAGGCGTTACTGTTGATTCAGATAATGTGATTATTCATTTAGATAGAAATGGTGCAGTAACAAGTGTTACAGGGAAATCCA
GTTGATCAAGTAGTGAAAATTCAATCGGTTGATCGGTTGATGCAAGTCAAGGAGGTTAAAAAAATTATTGCTTCTGATAATCCG
GAAACTAAAGATCTTGTCTTTTTAGCTATTGACAAACGTGTAAATAATGAGGGGCAATTATTTTATAAAGTCAGAGTAACT
TCTTCGCCAACTGGTGACCCCGTATCATTGGTTTATAAAGTGAACGCTACAGATGGAACAATTATGGAAAAACAAGATTTA
ACGGAACATGTCGGTAGTGAAGTAACGTTAAAAAACTCTTTTCAAGTAGCGTTTAATGTGCCAGTTGAAAAAAGCAATACG
GGAATTGCTTTACACGGAACGGATAACACAGGGGTTTACCATGCAGTAGTTGATGGCAAAAATAATTATTCTATTATTCAA
GCACCATCACTAGTAGCATTAAATCAGAATGCTGTTGACGCCTATACGCATGGAAAATTTGTGAAAACATATTATGAAGAT
CATTTCCAACGACACAGTATTGATGATCGAGGGGATGCCCATCTTGTCAGTTGTTGATGAACAACATCCAGATGCTTATGAC
AATGCTTTTTGGGATGGAAAAGCAATGCGTTATGGTGAAACAAGTACACCAACAGGGAAAACGTATGCTTCCTCTTTAGAT
GTAGTTGGTCATGAAATGACACATGGTGTGAACGGAACATACTGCCGGTTTAGAATATTTAGGACAATCAGGTGCGTTGAAT
GAATCTTATTCTGATTTGATGGGTTATATTATTTCGGGTGCATCTAATCCAGAAATTGGTGCGGATACTCAGAGTGTTGAC
CGAAAAACAGGTATTCGAAATTTACAAACGCCAAGTAAACACGGACAACCAGAAACCATGGCTCAATACGACGATCGAGCA
CGGTATAAAGGAACGCCTTATTATGATCAAGGCGGTGTTCATTATAATAGTGGAATTATTAATCGGATTGGTTACACCATT
ATCCAGAACTTAGGCATTGAAAAAGCACAGACTATTTTCTACAGCGTCGTTAGTAAATTACTTAACACCTAAAGCACAATTC
AGTGTAGCTCGTGATGCGATGCTTGCTGCTGCAAAAGTTCAATATGGCGATGAAGCAGCTTCAGTGGTATCAGCAGCCTTT
AACTCTGCTGGAATCGGAGCTAAAGAGGACATTCAGGTAAACCAACCAAGTGAATCTGTTCTGGTCAATGAA
```

EF1850
Seq ID 396
```
GTGCAACAAAACAACAATGAGATATCTACACAACACCAATTTGATGCCTACTGTAAAAAAGTATTACGCAATGAAGCTAAG
TCTATTCGAAAACGAAACGATAAAATAAAAGAAAATGAGGAACCATTAAATGATTTAAATGAAGGTAAGTATTCACAAACT
CATTTAGATGAGCAAGATGTTTATTTCTTATTTGGAATGGAGATATTAATTTCAGATCAGAAGTTATCAGCAGCAATTGAT
CAGTTATCGGATACAAGGAAAAAGATTGTTTTACTTTATTATTTTGCAGGTTTTAATGATACTGAGATCGGTAAAATTTTC
AACATGAGTACAAGTGGTATCTGGTATCAACGAACGAAAGCGGTTGAACAATTGAAAATGGAATATGGCTTATGG
```

EF1877
Seq ID 397
ATGAAAAAGAAACTTTTTCGGGATTGTGGGAATCGTTTTTATTAGCATACTCACTCTGATAGTTCTATTGAGTTTAGTTGGT
ACGGTAGCCGAGGCGACAGGGCTAGTTGATGATACAGTTGAATCAGGAAATCTATATTCCAAATATTCACTGAACAACTAT
CAACTAGATTTCTTTGTGGATAGCTCTTGGGATTGGTTGCCTTGGAATTGGGGTGATGGCTTAGGGAAAAGCGTGATGTAT
GGTCTTTATGCCATTACTAATTTCATTTGGACGGTGAGTTTGTATTTGTCGAACGCCACGGGTTATGTGGTTCAAGAAGCC
TATAAATTAGATTTCATTTCGGATACGGCTGAAAGTATTGGAAAGAATATCCAAACATTGGCTGATATTACTGAAAATGGA
CTCCAGACCTCCGGATTTTATTTTGGGTTTCTATTATTGATGATTTTAGCTCTGGGGGTGTATGTGGCGTATACCGGATTA
TTAAAACGAGAAACCACGAAAGCAGTTCGAGCCGTGATTAATTTTGTGATGATTTTTCTACTATCGGGTTCGTTCATTGCT
TATGCGCCTACCTATATCACAAAAATCAATGACTTTAGCTCAGATGTTAGTGAAGCGGCCCTTACCCTTGGAACCGAGATC
GTTGTACCCAATTCAGAAAGCCAAGGAAAAGATAGTGTGGATTTGATTCGGGATAGTTTGTTTTCGATTCAAGTCCAACAA
CCTTGGTTACTACTGCAATTTGATGATTCCAACATAGAAGAAATTGGTGAAGACCGGGTCAACAAGATTCTATCTGTGAGT
CCAGATGAAAACAAAGGCAAAGATCGGGAAGAAGCGGTCAAGGCAGAGATTGAAGACAATGACAATGCGAATCTCAGTATT
ACCAAGACCATGAACCGTTTAGGGATCGTGGTCTTTTTGGTATTGTTCAATATTGGCATTTCCTTCTTTGTTTTTCTTCTA
ACAGGGATTATGTTGTTCTCGCAAATTCTCTTCATTATCTTTGCGATGTTTCTGCCAATTAGTTTCTTATTGTCCATGTTG
CCGACTTATGAGAGTTTGGGCAAGATGGCGATTATTCGTTTATTCAATACGATTATGATGCGAGCGGGTGTCACCTTGGTG
ATTACGACTGCATTCAGTATTTCTACGATGTTTTTCAATATTTCGGCTACCTATCCCTTCTTTATGGTAGCTTTTCTACAA
ATTGTGACCTTTGCCGGTATCTATTTCAAATTGGGTGACATTATGAGTATGTTCAACCTTCAGAGTAATGACAGTCAGTCA
ATGGGCAGACGGGTCATGCGAAGACCGCAAATGTTGATGAACCGTAAATTGCGGCAACTCAATCGAAATGTGGGCCGAACT
CTGGCTTTTGGCGGTGCTGCAGCAGTTGGGAATAAGTTGGCAAAGGAACAATCAAAACCCAAATTTAAGCCAGCAGGTTCT
TCTCTACGAAAGAATTCACGATTACCTAACGATCGCGAGGTATCCTCAGACTCGGCTAAGGAGAATCCAATTTCAAATAAT
AAGAAACAATCCCGAATGAATTTGACTGGGAGAAAAATAGGCAAAGTTCTGGATACCCAGGCTTTGGTGAAAGATAAGGCA
AAACAAGTAAAGGATCAGGTACGGAATACGCCGACCAACTTGAAATACATCCTCCATAAAGGTCTGGAAAAGACGAAAAAA
GCACCGAAAGAATTTAAGCGAGGGCTTGTTCAAGAAAAAGCAGATCGAGAGAAGCTGCGAGACAAACAACGACAACGCAGA
GATGAGAAAATGGATGAAAAGCGAAAAAACTTGGGCGAAGTAACAGACCGTCATGGAAAAAGAAGAAGTAATGTTTCGATA
AAAGAGGATCCAAAACCTCAGAAGGATAGGATACTGAAAAATGAACTTCCAAAACGGATTGTGATGGTAAAACCGAATTCT
GAGGTAAGGAGAAGACTTGTTAAACAGAATTTGGTTGCGAGAGAGAATAACCAACTATTAAGTAAGAAAAATTCTTTCCAG
CAAGTTAAGCAACGATCTACTCTTCCGAAGAGAGCCAATCAAAAAGTACAAAGAATAATGAAGGCTCGTCCAAAACCAAAG
TCGGGTGATAAAAAA

EF2174
Seq ID 398
ATGAAAAAAATCATTTCAGGTATGTTAATTTGTACTGTATTATTAAATAGTTTTAGTGTGATTGCATCAGGAGAAGAGCTT
GTAAAAACAGAAACAACGGGGGAGACTGGCTTGGTCACTCAAGTATCACAAACAACGACTAATTCAACGGAGGATACA
TCTTCTGTTACAGAAGAAAATACCTCTGAAAGAGATAGTTCTACTTCAAGCACAAAAGAAGAGTCTCTTGATTCTTCAACA
AATAGTACAACGGAAGAACATTCTTCTGTTTCAGAAACAAACACAACAGATAGTAAAACAGAAGCATCCACAAAGCTCTGAA
GTAGAAAAAAAGACCATTGACCAAGATGAAGCGGATTATCAAGAAGCAGCACGTGAAGGAACGAACCATAAAAAAGGAACA
TATGCTATGAAAAACGGCCTTTCTTCTAGGGTAGCGCGAGCAACAGTTGCGAACGTCTATGCGAATGATCCTAATTTGCCT
GGCAAGAATTTTATTGACGTTTCTAGTTGGAATGGAGACATTTCAGTAGCTGAATATCAAAAAATTAAAAAGCTATGGTGTA
ACAGGTGTTTCGGTTAAATTGACAGAAGGAACTTGGTATGTTAATCCATATGCTGCGGGCCAAATTAGAAATGCCAAAGCT
GCGGGCTTAAAAGTTTCAGCATATCACTATAGTATGTATGTCTCTGCTGCAGCACAAGATGAAGCAAGATATTTTGCT
CAGGCGGCGGCTAACTCTGGCTTAGATAAAAATACCATAATGTTTAATGATGCGGAAGATCCAACGCTTACAAACAATGGG
CGAAACGCCCATGCTAATTCTGTTGCGTTTAACCAGCAATTAAAGGCATTAGGGTATAAAAATGATGCGTTGTATGTTGGC
AAATGGTGGTTAACAAATGGCTATATCGACACCTCCGCCTTCGGAAGAGACAGAGTATGGGTAGCGCAATATCCATATACT
CCAGATAGTTCCATGCAGTGGAATAATGACCATGGTGCTTGGCAATGGTCTTCGCAAATGTATTTCCCTGGTTTAGCTAAT
TATGAAGGGCGTCCGTTTGATATTTCCATGACCTATAGTAATTTTTTTAAATATGGGTAATAGTAGTGGGCCTGACTTAAGT
AAGTATTACACAACTAATCCAGGTCGAGTGATCATGAAAAATGACGATACTTTTTACCAAGATGTAGCGTTTAGAACTCCT
GGTTGGCGAGTTAAAAAAAAATACACTAGTTACCATTAAGGGGGATTGAATATAGTTCAGCAGGAATTCCTCGATTGGTCACA
GACCAAGGTTATTTAACAGCAAATAAAGATTATGTATTAGCGGCACAAAGCAATATTGATTTGTATTTTACGACAAATCCT
AAAAAAGTTCGTTTGAAAAGCGATGATTATTTCTATGCTGATCCAGAATTTAAGCAAAGGTTGTCTAAGGTTTCGAAGGGA
ACAATAGTAGAAGTTGAAGATTTAGCTTATACTCAAAGCGGTATATTTCGATTGAAAACAGCGAAAGGATATTTGACTGCA
AATAAAAATATAGTTGAACAATATAAAGGATTAGAAGATACTTACTACACATCTAATCCTGGACAAGTAATTACTAGAAAT
GAAGATACCTATTACAAAGATGTAGAATTCAAGACAAAAGCAAATAAAGTATCTGCAGGCAGTGTTTTAAAAGTAACTGCT
ATTGAGAAACAAAATCAGGAATTCCTAGATTGAAGATGGCAAATGGGTATTATTTCACGGCTAACAAGAACTATGTAGTA
GCTACTGGATCTTGGATTGCTAACTATCATACGGTTAATCCAGGGCAAATTATTATGAAAAATAGTGATAATTTTTATGGG
GATTCAGACTTCTTATATAAAGGAGCAGCTGTTTCCAAGGGAAGTTTAGTTCCTGTTGTGGGAATCGAGTACCGTGAAAAT
CAAGTTCCTCGTTTGATTACCCAAAATGGTTATTTAACGGCTAATAAAAGTTATGCACAAAAAGTTGTGCCAAACATCAAA
GATTATCTATATGATTATCCAGAGTATGTAGTTATGAAAACAAATGATTATTATTATCAAGATGTAAATTTCAGTAAAAAA
GGTGAGTTTGTTTCTAAAGATACATTAATCAAAGTTTTAAGTGTTGATTATACCGAAAACGGATTACCTAGATTAAAGACT
GCCAAAGGTTATATAACTGCTAACAAATCATATGTTACAAAATTGGTATCAAATTCAGATAATTACTTTTACTGAAAATCCG
CACCAAATTATTATGCGAGTGTCAGATAAACTATATACTGATGTAGAGTTTCGTAATGGTAGCAGAACCTTATCATCTGGG
ACTGTTGTTCCAGTTAAGGGCATTGAGTATTCCTCTAAAGGTGTTCCTCGCTTGAAGACAGAAGGAGGATATCTTACTGCA
AATAAAAAACTATGTGACAGCTGCTGGAAATACCAATAACAATTACTTTATCACTAATCCAAAAAAAGTTAAGCTCTTAATT
GATGATTGTTTCTACAATAATACTGAGTTTACTAAAAAAGGACAGGCTGTTAAGAAGAATACGATTGTTGAAGTTGAAGCA
ATCGAATATACAAACAATGGTATACCAAGGTTAAAGACGAAACAAGGGTATTTAACTGCTAATAAGTGGTATGTAGCAAAA
GTGGGT

EF2224
Seq ID 399
ATGAATAAAGCAGTTAAAAATTTTGTGAGTTACCTTATGATTACCATGCTGTTTATTTTAAATTTGTTACCAATGATGAAC
GCATTTGCTCAAGAAGTAACAAGTGATGCTGAGAAGACTGTTGAAAAAGACGGGCTCAAAGTAATAGGGGAAAATAGAAGAT
ACATCGTCTCAAGAAGATATTAAAACGGTCACTTACGAAGTAACAAATACGCGAGATGTACCAATCAAAGATCTCATTTTA
AAACAAAAAAATACCAATGATAGTCCAATCAAATTTGTTTTAGATACTTTAAGTGAAGAGCGAGGACCTACGTCTTTAGAA
GAACAAGCAAAGGTTGAAACAAATGAGAAAGATCAAACAACCGATATTAAACTGCTTAACCTTCAACCAAATTCGACAAGA
AAGATTACCATAAATGGGCAAATAACTACAAAAGCGTCGAACAAGTTATTGGTCAGCGTACTAATAGAAGATAATGAAAAA
GGAACACTAGTAATTGATTTACCTAGTAAAGATATACTAGCAGATAAAGAGTCGGTTTCGAAGGAAAAACAGGAAACTTCA
GAAACAAAAGTAGAAAATCAGGCAAACGAAACAGCCTCTTCTACAAATGAAATGACCGCAACTACTAGTAACGAAACAAAG
CCTGAAGCAGGAAAGGCGATAGAAAGTATTCAAGACAGCACTAACGCAGGCTACCGAAAGTCCTGAGCAACCTCCATTA
AAGGCGCAACCTACTGGTCCGTTAGTGCCACCAACACCTGGTCGGGGGTTTAATACACCGATTTATCAAAGCGTTCATAAA
GGAGAACTTTTTTCAACGGGAAATACCAACTTAAAAATTGCGAATGAAAATACGGCGGCTGCACAAACATTTTTAAACACA
CGAGGAGCAAGTAGTGGTTATGCAATTAACAATTTTCCTTTAGAATTTGCAGATGTTGACAATGATCCAAATACGTATAAC
TCTAGTCGGGCTTATATTGATTTAAATGGTGCAAAAGAGATTGCTTGGGCGGGCTTATTTTGGAGTGCATCTAGATATAAA
GGCCCTGCTTACGGAACAAATCTTTCTGATGAAGAAATTAGTGCACCAGTTCAATTTACTACACCAAATGGAACCGTACAG
CGTGTTTCGCCCCAAAGGTACCATCGTATCGATCAAGATGCAACAAACCCAGGACAACGTTTCGGGTACAATAACACTGGA
TTTTCTAATTATGCAGATGTAACTTCAATTTTACAAGGGGATAAAAGTGCGACAGGGAGTTATACGTTGGCAGATATTCCT
ATGACAAGTAGTTTAAATGGTCAATATCAATATTATAACTTTAGTGGTTGGAGTTTGTTTGTTGTTACAAAGGATCAGGCA
AGTAAGTCAAGAGCTTTTAGTATTTACTATGGAGCACGTGGTAATGCTGCTGGAACCAATAATGAATTTACTATGAGCAAC
TTTTTAACAGCAAAACAAGGAAATCTTGATCCAATTGTGACCTGGTTTACTGTTCAAGGAGATAAATACTGGACTGGAGAC
AACGCACAAATTAAAAATAGCGCAGGAACTTGGGTAAATATTTCGAACACGCTCAATCCAGTTAACAATGCTATGAACGCA
ACTGTGACAGATAACGATGAACATATGGTAGACAAGTATCCAGGGAAATTTGCGCCGGATCATCCTAATTTTTTAGATATT.
GATATCGATCGTATGGCAATTCCCGAAGGTGTTTTAAATGCTGGACAAAACCAAATTAATTTTAGAACAACAAGTAGTGGT
GACGATTATTCTACGAATGCGATTGGTTTTGCTGTAAACGCAGAAACGCCTGAATTTGAAATTAAGAAAGAAATTGTGGAA
CCAAAAGAAACTTACAAAGTTGGTGAAACAATCACGTATCGTGTCTCACTAAAAAATACAAAAGCTGATTCTGAAGCGATA
AATTCAGTCTCGAAGGATGCACTAGATGGTCGCTTAAATTATTTGCCAGGCTCTTTAAAAAATTATTAGTGGACCAAATTCT
GGTGAAAAAACGGATGCTTCAGGCGATGATCAAGCTGAATATGATGAAACCAATAAACAAATTATAGTCCGTGTCGGCAAT
GGTGCAACTGCTACACAAGGAGGTAGCTACAAAGCTGATACGGAATCAGTTGACAGGAATCAATTTAAAGCGCGGATAAATGAA
CGAGCAAAAGCAAATGAATTAGTGCCTAATAGCGCAACCGTTGAGGCAGTAGATATTTTAACATCTGCAAAAGTCAACGAA
ACTTCTAATATAGTAGAAGCAAAAAATTGCGGACGAACAAGTAACTGGCAAATTAACCGCAACAAAAACCGTTAATAACGCC
AAACCAAAACTAGGGGAAGAAATCGAATACACCATCAGCTTCCGCAACACTATCGAAAACGGCATCCTAAACAAAGTAGTT
ATCACAGACCAACTGCCAAAAGGACTGACATATGTCAAAGATAGCCTAACAAGTGTTGGTGATGAACCAAAGCCAACCAGC
TTGAAAGAAACCAACGGCACAATCACAGCGGAATACCCAAGTATCACCGATATGAAAGAACGGAGTATTCGCTTCAAAGTG
ATCGTTAACGAAGAAGCCAAAGCGGGCGAAACAATTCTCAACAAAGCCAAAGTAGATGACACAGTTAATCACCAGAAGAA
CCAGAGGTGCCTGTGGTACCAGAAACGAACGCAGGCAAACTAGCCGCAACAAAAACCGTTAATAACGCTAAACCAAAACTA
GGCGAAACTATCGAGTACACAATTAGCTTCCGCAATACCATCGAAAACGGCGTCCTAAACAAAGTAGTTATCACAGACCAA
CTACCAAAAGGACTGACGTATGTCAAAGATAGCCTAACAAGTGTCGGTGATGAACCAAAACCAACCAGCTTGAAAGAAATC
AACGGCACAATCACAGCGGAATACCCAAGTATCACTGATCAAAAGAGCGGAGTATTCGTTTCAAAGTGATCGTTAACGAA
GAAGCTAAGGCAGGCGAAACAATTCTCAACAAAGCCAAAGTAGATGACACAGTTAATCCGCCCGAAGAACCAGAGGTGCCT
GTGGTACCAGAAACGAACGCGTAGGCAAATTAACCGCAACAAAAACCGTTAATAACGCCAAAACCAAAATTAGGGGAAGAAATC
GAGTACACAATCAGCTTCCGCAACACCATCGAAAACGGCGTCCTAAACAAAGTAGTTATCACAGACCAACTACCAAAAGGG
CTGACGTATGTCAAAGATAGCCTAACAAGTGTCGGTGATGAACCAAAACCAACCAGCTTGAAAGAAATCAACGGCACAATC
ACAGCGGAATACCCAAGTATCACCGATACAAAAGAGCGGAGTATTCGTTTCAAAGTGATCGTTAACGAAGAAGCCAAGGCA
GGCGAAACAATTCTCAACAAAGCCAAAGTAGATGACACAGTTAATCCGCCCGAAGAACCAGAGGTGCCGGTGGTACCAGAA
GCGAAAGAAGGCAAATTAACAGCAACAAAACCGTTAATAACGCTAAACCAAAACTAGGGGAAGCGATCGAGTACACAATC
AGTTTCCGCAATACCATCGAAAACGGCGTCCTAAACAAAGTAGTTATCACAGACCAACTACCAAAAGGACTGACGTATGTC
AAAGATAGCCTAACAAGTGTCGGTGATGAACCAAAACCAACCAGCTTGACAGAAGCCAATGGCACAATCACAGCAGAATAC
CCAAGTATCACCGATATGAAAGAACGGAGTATTCGTTTCAAAGTGATCGTTAACGAAGAAGCGAAAGCGGGCGAAACAATT
CTCAACAAAGCCAAAGTAGATGACACAGTTAATCCGCCCGAAGAACCAGAGGTGCCGGTGGTACCAGAAGCGAAAGAAGGC
AAACTAGCCGCAACAAAAACCGTTAATAACGCCAAACCAAAACTAGGCGAAACTATCGAGTACACAATTAGCTTCCGCAAC
ACCATCGAAAATGGTGTCCTAAACAAAGTAGTTATCACAGACCAACTACCAAAAGGACTGACGTATGTCAAAGATAGCCTA
ACAAGTGTCGGTGATGAACCAAAACCAACCAGCTTGAAAGAAGCCAACGGCACAATCACAGCGGAATACCCAAGTATCACC
GATACGAAAGAACGGAGTATTCGTTTCAAAGTGATCGTTAACGAAGCGAAAGCGGGCGAAACAATTCTCAACAAAGCC
AAAGTAGGTGATGGCATCAATCCACCAGAAGAACCAGAAGTGCCCATTACGCCTGAAGAACCAGCGAAAAATAAAAAGGAA
ACAAACAAAGTTGTTACAGACCAGAACAAACCTACAAAAAATAGTAAAAATGAAATTGCCATAAACAAGAAAGAAACGTCT
AAATCATCCTATTTACCAAAAACAGGAGAAAAAGTTCAAAAAATATTTGCGTATCTGGGTGTAGGACTGATCTTAATTGTA
TTAATTCTCTATGTTATAAAGAGAAACAAAGAAAAAGAGGAG

EF2318
Seq ID 400
GTGAGCCGGGAACGGTTTCAGGCACGGGAGAAAAAGGTTCAGAAGATGGGCCGGGACGGTCTGGTGGAACAGAACCGCGCC
ACCGGCGAAGAAAAGCGCGTCAGCCAGCGCACGGCAGACACCGCCTTTGACCGCGCCCGCCCTACGGAACGGGAGGACATC
CGCCGCCCGGCAGCACGGGGCGCAGACACCGGCAAAAAGCGGAAACAGCCCCGCACTGCGCCGGAGATGGCGGCAGAACAG
ACGGAACCGCTGCTTACGCCGGAATATCTGGCAGACAGTACTGCCGAACCGCCGCCCGATGCCCCGGCATCCATGCGCGGC
GCGGTGGATATGCCGTTACTGGATGCGCCCATAGCCGAAGAACCCACTCCCGCGCCGGTGAGGAAACAGCCCCAAAAGCGG
GGGAAAAAGGGCAAGAAGAAGCAAAGCACGAAATTCACGGAAGACGCTGCCCGACCGGGTAGTGAGGAAAAAACGGCTATG
CGCGGTGTGGAGGGCGCTCCTGCCCCAAAGCGGCGCACTTCTGAGGGCGGCTCCCGCCTGCAATTTGACAAGGAAGAAGAA
CCGGCAGACCCCGCAGCCGACCGCGCAGCAATCAGAAAGCGGCAGACAGCAAAGCTGGCCGAACACGCCGCAAACCGGAG
GATAAGCTAAAATCCGATGGCAATCCAAAGGCCGGGGAAACACGCCTGAAATTTGAGGACACTCCCAAAGATACGGACATT
TCCGGCGATACTGCGCCCGATATGCCCACCCGCCAACAAAAGAAATATAACAAGGCCGCCTGCCGCACGGAACGCGCAGAC

CGCCGCGTGGAGCAGGCACAAATGAAGCTGCCTGCCAAACGCCGCCTGCACATGGATCACCAGCCTGACGAAGCCAGCGGG
AAGATAAAGCGGCGGCTCCGTTTTGAGGATGAAATCCTGCCGGAGTACCAAAAGCCCTCCCTGTCCTCCCGTGCCGGAAAT
GCGGCAAAAACGGCGGCAGTCTTAAAGCTGCATGGCAAACTCCGTGAATATGAGCGTGACAACGTGGCTCTGGAATCTGCC
CACAAAATCGAACTGGTTGCCGAACAGGGCGCGGTCCTGCCGCTGGGAACGGAACCGCCGCCGCTCCAAACCCTAC
CGCACTTTGCGGAAAGCCCAGCAGAAAGCGGCGCGGGCGCACACTGACCTTGCATGGCAGACCGCCCTGCGGGATAACCCC
GAACTGCAACGCAAAAACGCGCTGGCAAAATGGGTGCAGAAGCAGAAAATCAAGCGGAAATATGCCCAGGCCGCCCATGAA
GCCAAACAGAGCGCGAAATTTACCCAAAACGTCCTGACCGCCACCGGCAAGATTGCCCGCGCCATCGCCCAGTATGCGGCG
GCACATAAGGCGGTCTTTCTGGCGGTTGCAATGCTGGCGCTGGTGGTGATGTTCTTCGCCACGGGCCTGACCTCCTGCACC
GCCATGCTGTCAGGCTTCCAGTCCTCGTACATTTCCGCGTCTTACATGGCTAATGAGCAGGAAATCTGCCAGAGCGACCTT
TATTACACGGAAAAGGAAACGGATTTGCAGATTGATATTGACAGAAAATTACCCCGGCTATGATGAATACCGC
TACAACATCGGTGAAATCAGCCATAACCCTTATGAACTGTTAGGCTACCTCTCCACGGCGTTCAATGCTTTTACCTTTGCA
GAAGTCCAGCCCGAAATTGACCGCATTTTTTCCCGGCAGTACACCCTCACCCGTGAGGTGATCGTGGAAACCCGGTATGAC
GATGACGGCGACCCTTATGACTGGTATGTTTTGCAGACCACATTGGCCGTGCGCCCGCTGTCCTCTGTGTTACAGAGCAGC
CTCACCCCCGGCGAACAGACCGACCGCTACGGCGTGTATATGCAGACCTACGGCAACCGGCAGGCGTTTGGAAATCCCTTT
GGTTTTTCCTGGCTTGGATATGTGAGCAGCGGCTACGGCTGGCGGGTACACCCTGTAAACGGCGAAAAGAGCCTGCACCGG
GGCATTGACCTCGCTGTGGCGCAGGGGACTCCGATTCTGGCGGCGCAGGACGGGCGTGTGGTATCGGCCGGCGATGCAGGC
AGCTATGGGCTGTGCGTGGTGATCGAGGACGATAAAGGCTACCAGTCCCGGTATGCCCACTGTTCCAGCTTAAACGTCAGC
GCCGGCCAGGAAGTAAAACGCGGGGATGTGGATTGCTGCGTGTCGGCAGCACCGGCAACAGCACGGGGCCGCATCTCCATCTG
GAAGTCATGTTAAACGGCGAATACTTGAACCCTTATTTCTTCGTGGACAATGGCGACGATGGAACCGGCGCAATCCCCGGC
ACCCCCGGCGGGCCGGCTATCCCGGACTATTCCGGCGAACCGATGGGCGATGGCAGCTTTGAAGCCATGCTGCGGGAAGCG
GAGAAATACCTGGGCTATCCCTATGTGTGGGGCGGCTCCCATCCATCCACTTCCTTTGACTGTTCCGGTTATGTTTCGTGG
GTCATTAACCAGTCCGGCGTGGGCAGCGTCGGGCGGCAGACGGCACAGGGCCTTTACAACCTGTGTACGCCGGTATCTGCG
GTGAACGCACAGCCCGGCGACCTGATTTTCTTTACCGGGACATACAGCTCGCCGGGGCCGGTCAGCCACGTTGGGATTTAT
GTGGGAGGCGGGCGGTTTATTCACTGCGGCGACCCGATTTCCTACGCCAATACGGGAAGCCCCTATTGGTCTGCCCATCTG
TACGGGTATGGCAGAATACCC

EF2704
Seq ID 401
ATGACGAATAAAGAGCCTTGGGAGGCGTGGAGTCCCGCAAAAGTTAGTTCATTCCAGGAAGATTTTTTAGCTTGGTATGAA
CGAGAAAAACGCAATTTACCTTGGCGAGCGAATACAGATGCATATCGTATTTGGATTTCTGAAATTATGCTACAACAAACT
CGCGTAGATACAGTCATTGATTATTTTTATCGATTTATGGAATGGTTTCCGACGATTCAAGATTTAGCGGAAGCGCCAGAT
GATAAGTTGTTGAAAGCTTGGGAAGGGTTAGGTTACTATTCACGAGCGCGTAATTTAAAAGTGGCAGCGCAACAGATTGTT
TCAGAATTTGGTGGGAAAATGCCTGACACAATCGAAGATATTCGGAGTTTAAAAGGAATCGGCCCCTATACGGCTGGTGCG
ATTGGCAGTATTGCCTTTAATCTCCGCGAACCAGCAATTGACGGTAACGTCATGCGGGTAGTGAGCCGTTTATTTGAAATT
GATGCAGATATTGCAAAAGCGAGTAGTCGTAAAGTATTTGAACGGCGATGCTTAAAATCATCGATCGTGAGCGTCCAGGG
GATTTTAATCAAGCTTTAATGGATCTAGGCTCAGCCGTCTGCACCCCAACTTCCCCTAAGTGTGAAAGTTGTCCGCTTCAG
CAATATTGTGCGGCCTATCAGGCAGATAAAATGACCGCTTATCCAGTCAAATCAAAAAAAGTTAAGCCGAAAGATGTGTAT
TATGTCGGCACTATTATTGAAAATAAAAAGCAAGAATTTTTATTAGAACAACGCCCTGAAACCGGTTTGTTAGCCAATATG
TGGCTGTTTCCGATTGAAGAAATTAGCAAAAAACAATTTCAACAATTGCAAAAACTGGCGCAACCAGCAGAAACGGAAAAA
CAATTAACTTTAGAGTTAGAGCCTGTAACAGAGCCTTTGGTCGCTGAAGAACCTGTCAACTTTTTTACGGACTATGAAACA
GTTGTCTGGCAAAAACGTACCTTAGGCGAAGTCGTTCATATTTTCAGTCATTTAAAATGGCATATCTTAGTCTTTTATGGC
CGAAATACAGGCGAATTAGCCACACTAGAGTCGCAGCGATGGGTGGCAGCACAACAATTTTCAGACTATGTTTTTCCAAAA
CCACAGCAAAAAATGGTGGAATTGTTCAAAAAAGAACACGAAAAAAAA

EF2713
Seq ID 402
GTGGAAAAGTCAACAGTTACTTCTTTAACTCAAGAAACTTCAGCAACAACAATTAATGCTTCAACTGATTCAACAGCACTT
ACAGCTGAAAGCGAGGAACTTCCATCGCTTCGTCAAACACTTCTTAATTACGTAGGTATGTATGGTTTAACAGAAACCTTG
ATAAATCGTTTATCTGATGACGAACTGGACTACGCAAAAAAAGTTTCGTTTCATTTTGTAAACCAAGACATCAGTGGAACA
GCGAGAATGATTACTAAATTATATGGTGAGAAACCTATCCCGGAGGACTCCTACTCTACGATTACTCAACATTAACGATT
GATGACTTAAAAAATTATTTACCTCAAATTCGACTTTCGTTAATTTATGTATATGATTTAAATAGTGATGTTGTAAATAAT
TTAAGCGATCAAACACTCGTCGATTTAATTAATCAAGTCAAGGTAGACTACGCTAACCAAAATTACCCATCTGATGTCCGT
GGTGATTATGGTTTAGCAGCAATGGCAGATAAAATAAAAGCAAACGATTATACGTCAATTAATCAAAGTGCTGAAAGTGTT
TCTTCGGATACTACTAATACAGAATCTACATTACAAACAACTACGAGCAGTTCGAAGAAAGCTACTACATCAAGTTCAACC
GAGCACAAAAAGGGGATATTTCCCAGTACTGGCGAAAAAAAATCAGTTTTATTTACTATCATTGGAATCATCTTACTATCT
TTAGTTAGTATATTCATTATAAAAAATAAAAAGAAA

EF2802
Seq ID 403
ATGAAAAAGAAAATTTTAGTAGGAGCGCTTGTCGCTCTATTTTTTATGCCTTTAAATGTATTTGCTGCAAAAGGTGACCAA
GGCGTTGACTGGGCGATTTATCAAGGCGAACAGGGTCGTTTTGGTTATGCACATGATAAATTCGCCATTGCTCAGATTGGC
GGCTACAATGCTAGCGGTATTTATGAACAATACACATATAAAACGCAAGTAGCAAGTGCTATTGCCCAAGGTAAACGTGCG
CATACCTATATTTGGTATGACACTTGGGGAAACATGGACATTGCGAAAACAACAATGGATTACTTTTTGCCACGTATTCAA
ACGCCTAAAAATTCCATCGTTGCATTAGATTTTGAACATGGAGCGTTGGCTAGTGTTCCAGATGGATATGGAGGATATGTA
AGTTCAGATGCCGAAAAAGCAGCAAATACAGAGACAATTTTGTACGGTATGCGCAGAATAAAACAGGCTGGCTATACTCCA
ATGTATTACAGCTATAAGCCATTTACACTAAATCATGTAAACTATCAACAAATCATCAAAGAGTTTCCTAACTCTTTATGG
ATTGCTGCGTATCCTATCGATGGTGTGTCACCATATCCATTGTATGCTTATTTCCCAAGCATGGATGGTATTGGCATTTGG
CAATTCACATCCGCTTATATTGCAGGTGGTTTAGATGGTAACGTAGATTTAACAGGAATTACGGATAGTGGTTATACAGAT
ACAAATAAACCAGAAACGGATACGCCAGCAACAGATGCAGGCGAAGAAATTGAAAAAATACCTAATTCTGATGTTAAAGTT
GGCGATACCGTCAAAGTGAAATTTAGTGTTGATGCTTGGGCAACTGGCGAAGCTATTCCAGATTGGGTTAAAGGAAACAGT
TATAAAGTGCAAGAAGCAACTGGAAGCAGAGTATTGCTAGAAGGCATTTTGTCATGGATTAGCAAAGGTGATATTGAACTA

TTACCAGATGCAACAGTTGTCCCTGATAAGCAACCAGAAACAACACATATAGTTCAACATGGCGAAACATTATCAAGCATT
GCTTACCAATACGGTACCGATTATCAAACTTTAGCTTCGTTAAATGGATTGGCTAATCCAAACCTTATTTACCCTGGACAA
GTTTTGAAAGTAAATGGATCAGCAACAAGTAATGTTTACACAGTCCAGTTGGGTGATAATTTATCAAGTATAGCGGCTAAA
CTTGGCACGACTTATCAAACATTGGCAGCTTTAAACGAATTAGCAAATCCTAACTTGATTTATCCGGGTCAAACGTTGAAC
TAT

EF2813
Seq ID 404
GTGTCAAAACAAGAATCCGATGTTGTCTTAAATTTTAAGATGAATGGAGAAATAAACTATTCACGAACAATTAAAGACATT
AATAAAGAAATGAACTTAGCGGCTACCGAGTACAAAAACCAGGTATCCGCAATGGATAAAAATGCAACTCAAACTGAAAAA
TTAACAGCTACAAAGAAAAAATTAGAAAAACAATTATCTTTAGCTGAACAAAGAACAAAATTATTACGTGAGGAATACGAA
AAATCAGTAAAAGAAACTGGTGAATATTCAGAGCAATCACAAAAGCTGTATAAGCGTTTATTGGAATCCGAAACAGGTGAA
AATAAACTGCGTTCTGCATTGCAAAGTACCAATGAAGCTTTGAAAGAGCAAGGTAATTTATCAATAAAAACAGCTGAAAAA
CTAGCCAAAATTGAAAAAGCTGGAGACAAAATTAAATCAGTTGGGCAAAAACTGTCTGTTGGATTAACAGCACCAATTATG
GGAATTGGTGCTGCTTCTATTGCCGCATTCAAAGAATTAGATGAATGTTTGGATAATATAACAACAGCAACAGGGGCTACT
GGTAGTCAGCTAGAATCTTTACAAGCCAGTTTTAAAACAGTAGCAGGTCAAATACCCGCGGATATGCAAGATATATCAACT
GGTATTGGTGAAGTAAATACTCAATTTGGCTTAATGGATAAGCAATTGGAAGATACAACAGGCCGAATGCTTAAATTTTCA
GAAATTAATGGCTCAGATGTTTCTCAATCAACTATTAATGCAAAAAATCAATGGACCTTTTTAGGTTGTCTATTGAGGAT
TTGCCAATGATTTTAGATTCAGTATCTAAAACTAGCCAAGATACTGGAGTAGGGGGTAGATCAGTTATTTGATGCCGTAAAT
AGAGGTGCGCCCCAACTCAAAGCTATGGGACTTGGTTTTTCTGAATCAACTACGTTAATAGGTCAAATGGAAAAAGCTGGT
ATTGATTCAGCAGGAACTCTTGGCTATTTGGCAAAAGCTAGTGTCGTATATGCGAAAGATAATAAAACCATGCAAGAAGGG
CTTAGCGGAACTATTGAATCTATTAAAGGGGCCACAACTGAACAAGAAAAACTCACTATTGCTAGTGAAGTTTTTGGAACT
AAAGCTGCTTCAAAAAATGGTAGAAGCTATTGATAGCGGAGCGTTGTCAATGGATGGTTTAGCAGATTCAGCAAAAAACGCA
GCTGGCACTGTGGATCAGACGTTTAGTGATATTCTTGATCCAATTGACCAAGCAAAGTTGGCACAAAATCAATTTAAAATA
GCAATGGGTGAACTTGGAGAACAAGTACAAATAGCATTATTACCAGCATTTCAAGCTGCAACGGATGCAATAAAAAAAGTT
TCAGAATGGTTTGGAAGTTTAACAGATAGTCAAAAGCAAACCATACTGAAAATAGCTGGTGTTGTGGCTGCTATCGGTCCA
GTATTAGTAGTTTTAGGAACACTTGCTAGTTCCATTAGTAGTTTGATTCCAGTTATTGCTTTTATTGCGTCGCCAATTGGT
TTAGTAATTGCGGCGGTTGCCGCTTGGGTAGCTGCAATCGTAGTTGCATATAATAAAATCGGTTGGTTTAGGGATTTTATC
AATACCTCCTTTAAAGTAATTAAAGATATTGTGGTTGGTGTATTTAATGTTTTGAAAGATACGACAAAATCTACTTTTGAT
TTCATCACAGGATTTATTGGTGGTGCCATGGATGGGGCTGTCAAAATTATTGGCGATTATGTAAATGAAATTAAGCGTATT
TTTGGCGGTATCGTTGATTTTGTAACGGGAGTATTTACTGGAGACTGGTCAAGAGCGTGGCAAGGTGTTGTTGACATTTTT
GGTGGTATTTTTGAAGGCATCGCTGCAGTAGCTAAAGCTCCAATCAATGCCATGATTACGTTAATCAATGGATTTATTGGT
GGATTAAACAATATAAAAAATACCTAAATGGGTGCCAGGAATTGGCGGTAAAGGATTTCATATTGGAAAAATCCCTTATTTA
GCAGAAGGTGGAACTATTCTAAATGGCCAAGCCATTGTTGGTGAAGCTGGTCCTGAACTATTAACCGCTAAAAACGGCAAG
ACAACAGTAACTCCATTGTCACCAGAAGAAAAAGCTCGTGGAATTGGTGGTGCTTTGAAAGGTGGCAACACTATTGAGCAA
CATGTTCATATTGGCCAAGTAGATGCAAATAATCCGAGTGAGTTAGATCGAATGAATCGCAAGCTTTATAAAGCAAGTGCG
CAAGCTTTCTATGACTTAGGAGGTGTTCCAACG

EF2820
Seq ID 405
ATGAACAAAGAATTATTGCGTCAATTACAAGCTCGTCACGAGAAACGATTAAGTGATTTACAAGGCAAAATTGAATCTGGA
GAAGTCGTGAAGCAGATTTAGATTCAGTTAATGAAGAAATTGATGGTTTAATCGATGAATTAAAAGCCATTAAAGCTGAA
TTAGGGGATGATAATTCAGAATCTGGTGATGGTGAAGGCGATGACGGGACCGCTAAATCTGATAATACTGATGATGAAAGT
AAAGAAGATCGTGAGAAAGATACGAACGGAAACAACGATGATAAAAACGAAGAAAATCGTGGCGGCATGATTAGCCAAGAA
CAGCGTGATGGTTTGTTACGCACAATTCATGAAGGAATGGAGGCTAGAAATGTGATGTCTAATGAACAACGTGAAAAACAA
ATTCGTAAAGCATTTGCTGATTTTGTTATTGGTAATATTTCAGAAAGTGAAGCACGTGCATTAGGTATTGAAACAGGCAAT
GGTTCAGTGACAGTACCAGAAGTGATTGCATCCGAAGTGATTTCTTATGCTCAAGAAGAAAACCTATTGCGTAAATACGGA
ACTGTGATTCGCACGGCTGGCGATGTGAAGTATCCAATTCTTGTGAAAAAAGCAGAGGCTAATGTAACAAAAAAAGAACGT
ACGACAGATATTACTGCGAAACAGCGATTCAATTTGATGAAATTCTACTTGATCCAGCAGAATTTGATGCATTAGCAACTGTA
ACGAAAAAACTATTAAAAAATGTCTGGTGTGTGCCAGTAGAAGATATTGTTGTAGAAGAATTGAAAAAAGCATATGTTCGCAAA
GAAATTAATTATATGTTTAATGGCGACGATGCAGGAAATGAAAACCCAGGAGCTTTAGCTAAAAAAGCTGTTGCATTTGAA
AAACCTGTAGATTTAACAGCTGCAGGTGCTGGTCAAAAATTATATGATGCATTGATTGAATTTAAAAATACACCAGTAACA
GAAGTAATGAAAAAAGGTCGTTTTATTATTAATCGTGCAGCTTTAACTGCTATTGAAAAAAATGAAAACAGATGATGGATTC
CCATTGTTACGACCATTCACGCAAGCAGAAGGTGGTATTGGTTATCAATTAGTTGGTTATCCAGTTGATTGGACCGATGCA
GCAGATAAAAAAGGTGAACCAGATACACCGTATTATATTTTGGTGATTTTTCTGCTTTCAAAATTCAAGAAGTTATTGGA
GCGTTAGAAATTCAAAAATTGGTTGAAAAAATTCTCTGGAAAAAATCAAGTTGGTTTCCAAAATTTACAACTTGTTAGATGGC
CAATTAGTTTACTCACCATTCGAGCCGGCTGTTTATCGTTATGAAATTACAAAACCAGTAGGTGGT

EF3082
Seq ID 406
ATGAAAAAGAAATTTTTAGCAATGATGGCAGTTTCAATGATGGGACTGTTAATGTTAAGTGCTTGTCAAACAAATAAAAAA
ACAGCAGATTCTGCAACAACAGAAACAACAGCTAAAACGGAAGTCACAGTCAAAGACACCAATGGTCAATTAACCGTTCCC
AAAAAATCCTAAGAAAGTCGTTGTTTTTGATAATGGTTCCTTGGATACAATGGATGCACTAGGTGTCGGTGACCGCGTGGTA
GGTGCGCCAACTAAAAATATCCCTGCGTATTTGAAAAAATACCAAAAAGTTGAATCAGCAGGCGGCATTAAAGAACCAGAT
TTAGAAAAAATCAATCAACTAAAACCAGACTTAATTATTATTTCTGGTCGTCAACAAGATTATCAAGAACAATTAAAAGCC
ATTGCGCCAACCATTTACTTAGCTGTAGATGCCAAAAATCCTTGGGCATCAACGAAACAAAATATCGAAACGTTAGGCACT
ATTTTTGATAAAGAAGAGGTAGCTAAAGAAAAAAATAACTGGCTTAGAAAAAAGAAATTGCTGACGTGAAAAAACAAGCAGAA
GCTAGCGCGAATAATGCGCTTGTTGTGTTAGTTAACGAAGGACAACTTTCCGCTTACGGAAAAGGCTCTCGTTTCGGTTTA
ATTCATGATACATTTGGCTTCAAAGCAGCAGACGATAAGATTGAAGCTTCCACTCATGGGCAAAGTGTTTCTTACGAATAT
GTTTTAGAAAAAAAATCCTGGGATTCTCTTTGTGGTAGATCGCACCAAAGCAATTGGTGGCGACGATTCAAAAGATAACGTC

GCTGCAAACGAATTGATTCAAAAAACCGATGCTGGTAAAAATGATAAAGTCATTATGCTTCAACCAGATGTTTGGTATCTA
AGCGGTGGTGGATTAGAATCAATGCATTTGATGATAGAAGATGTTAAAAAAGGATTAGAG

EF3256
Seq ID 407
ATGAAAGTGAATAAATTTGTTAAAGGTTTTGCGGCAATTGCATTATCAAGTTTAGTATTAGCAGCATGTGGTGCAGACAAG
AAAGACAACACAACGAACTCTTCTAGCGCAGCATCTTCAGAAACGAAAAAATCAACGGAATCATCAGCACCAGCGAAAAAA
GTTGCCGGTGGCGATTTAAAAGATGGTACGTATAAATTAGAAGAAAAAAATGAAAAAAAATGGTTACCGTGCAGTCTTTGAA
ATGACTGTAAAAGACGGCAAAATCACTGAATCTAAATATGACAACATCAATGCTGACGGCAAGTCTAAAACAGAAGACACT
AAGTATGAAGAAAGCATGAAAGCAAATCTGGTGTTGGACCAAAAGAATACATCAAACAATTAAACGATTCTTTTGTTAAA
GCACAAAGCGCAAGCGGTGTGGAAGTAGTAACTGGTGCGACTCATTCATCTGAATCATTCCAAAACTACGCACAACAATTA
ATCCAAGCAGCACAAGCTGGTAACACAGACACAATCGAAATCGACAATGGGCAACATTGAAAGATGGTACGTACTCATTG
AAAGAAAAAAATGACTCAAACGGCTACCACACAACATTCTCAATGACTGTGAAAGATGGTAAAGTGACTGAATCTAACTAC
GATAACGTGAACGCTGACGGCAAATCTAAAAAAGATGACACTGAATACGAAAGCAAAATGAAAGACGTTGCTGGCGTTGGA
CCAAAAGAATATATCGAAACATTAAACAAAGAATTTGTTAAAGCAATGGGCGAAGAAGACGGATCACCTGCAGGTGTTGAA
GTAGTAACTGGTGCAACACACAGCACACATTCATTCATCAACTACGCACAACAATTAGTGAACGCTGCTGAAAAAGGCGAC
ACAACTGAAATCGTTGTTGACAACATCGTAACAAAA

EFA0021
Seq ID 408
ATGGACTTTCAAAATATACAAACTGTAAAAGAATTAAGATACGCACTGAAACAATACAAACAAACACTCGCAACTACGATT
TTTGATGATACACAAGGCTATATCAGTGTGTCGCCTTTAAAAAAGAAGTGATAGAAAAAGACGGAAAAAACGTTCAGTTTTTA
TGTTTTTATCCTAATTATGAGGTCGCTCTTTATCCAGTCGAAACGTTACTGGCTTACCTTGAAAAGTTACCTGAACAAGCA
ATCGTTACGATTAAAGACCATGACGGTAACTACTCGTCACTTTTTTCTAACAATTGACGAAACGATAGACCAAGGAAACAAC
TACCAATGGCTAGTGATTAGCGACAAAAAATTCTTACGAACGATTTTTTCTTTACAGCTCAAGAAAGCACGCAAGAACAAGCT
TTTGAGAAAATGTTAAAAGAACTAGAAACCGACCATGACAGTACACTAGAAGTTGTTCATAACTGGTTATGCCGACAAACA
GACGAAGCGCTTTTTTCTAGGTATTCTCAAAGAAGATAGAACGTTAAAAGGTGCTATGACTTACTGCATAGGAAAAGCCCAA
GAACAATCCAAAGACCAATCAAGCGCCATGGTGGCGGACGAAGTGGTTTTCAGTTGGATAAAAGAGTATTTCTTATTAGAG
AAATTACCAGAAACCAAAGCAGTTGGTAAAGTAACGACCAAAAAATCAGAAACAAAACCTGAAAGAAAAGAGGTGGACGAA
CAAATTGATTTGTTCGAGTGCGTA

EFA0044
Seq ID 409
ATGATTATTTTATCAGGACTTTTTATTTTAGGGATAGGAATAGTTGGAGGTTATCAGCTAGCCACGTTACCTAAATTGATT
GAAATGAAACAGCACAAAGCAATACAGAACCATTTCAATGTTAAAGGAAATGAGTACACCTATTATCAAGAAGACAGTGAA
AATTACATTTTATCCTTGGAAGATACAGAATATCGAATTAAATTTTCCAAGAACACCCCGTTAAAAGTGGTATTCACTGAA
ATTTTAGAACCTATG

EFA0052
Seq ID 410
ATGAAAAAGATCGCAAGTACAGGGTTAAGTATTTTAGTCGCAACAGGGGTAGCAGGTATTGGGGGAAATGAAGTACAGGCA
GCAGAACAAGCGCAACCAAAAACACCTGAAAACAGTTCTACAGAACAACCAACAGTGAAAGCTACACAAACAACGGAGCAA
GCTATTACTGAAAAACAGCAACAAGTAACAGAGAAACAAGCCAATTGTCGATCAAAAACAACAAGTCGCTGACACTGCGAAA
AAAGAAAAAGACGCCATTGATCAATCTGTTAAAGACCAACAAGCAGTGGTCGATCAAAACAAAGACGCATTGGATCAAAGT
CAACAAGCAGTGACTGACCAACAAGCAGTTGTAGACGAAGCTAAAAAAGTCGTGGATGAAGCAACACCTTCAGCCATTGAA
AAAGCCAAAGAGCAAGTGGCTACGGATACACAGGCAGTTGATGACCAACAAAAAGTAGTAGATCAAGCTCAAACAGACGTT
AACCAACAACAAGCTGTTGTCGAAGAAAAAGCAAAAGAAACGAATGCTGCTAAAGTACAAAATGATAAAGATCAACAAGCA
GTAACAGCTGCGAAACAAGAACAAGCCAAGCTTGAAGAATTAGCGAAAAATGCGGAACGCGGAAAAAGCAAAGGCAGAAAAA
GAACAAGCAGCAAAAGAAGCTGAACTGGCTAACAAACAGAAAGAAGAAGCCAAAGCTAAAGATCAAAAGACGAAGGACGAT
CAAGCAGTCGCAGACCAACAAAGCGTTGTGACAACTAGTCAAGAGAAAGTAGCAGACGCAAAAGCAGATACAGCTGCGAAA
CAGGCAGATTTAACTTCAAAAGAAAATGCTTTAAAAGACAAGCAAGCAGCAACAAAGCAAGCGCAGAATACTTTAGATAGC
TCAAAAGAAGAGTTAAAGGGGCATAAAGGACTTAACTTACCTGCTAACTTCACACCTGATTATTACAAAAAAATTAACTGAA
CAAGAAAAGCAAGCAATGGAAAAAGAAGCATTAGCATTAAATAAAGTTTTTTCCTGAAAATCAAGCAGATGCGGCAAAAGTA
ACGGAAATGATCAATTGCTCAAAAATCCTACCGAAAAAACAAAAGCAACAAATGAGCCGATTACGTTGTAGGACTTATCAATGAT
GTTCGCGAAAAAGCTTGGGTTACAAAAGTTGAAGATTTCTAACCAAGCTATGAAATTTGCTTGGGATGTAGCAAAATATGAT
AATCCCAAAGAATTTGATCATGACGTAAATGCGATCAATCGTGCAGCAAAAGAAAATGGTTTTAAAGAATTCCCTGGACAA
AACTTTTATGAAAACCTAAGTATGGGAAGATTTACGACACAAGAAGGTAAAGTTTCTATGTATGACTTTGAAAAAGCTGCT
CGAAATGCACTTGTAAGCATGTTGATGAACGATGGACATTCTGGCTATTCCCATTTAGATTCTTTATTAGATGCAAATGAA
ACAAACATGGCAGTTTCTATTTCAGGAGATTTAAATGATATTTCAGCAAAAATCCATATCATTAGCTATAATCAAACAAAA
TTAGTTGAAGCAAATACCTATGAAGAAGGCACTGCACCCAGTCGTTAAGAGTAAAGAAACCCTTCAAAAAGAAGTAGCGACC
AATCAAGAAAAATTAGCTACCGCACAACAAGCAGAATCAGACGCACAACAAGCAAGCTGGTGCAAGTCAGCAAGCCTTAAAT
ACAGCCAAAACAACACAAGCAACAGCAGAAAAAGAACTATCTGTTCATAAAGCGACATTGGCTAATCTTCAAGCAGTTGCG
ACTAAGAGTACGACAAACTATGAAGAAAAGTACGACAAACTGCGGCGGCAGAAACAAACCTTCAACAAACAAAGATCAA
TTGGCAACAATCAATGAGTTAATTCAGAATCGAGCTGTTGTTTTAGAAAAAGCGAAAACAAAAGTTGCAGACGCACAAGCG
ATTGAACAAACGTCTGCTAAAGTGCTGAAAGAAAAACAAGAAGCCCAAAAAGCAGAAGAAAACACATTGAATAGCTTGAAG
GAAGTATTGGATTTAGCAAAAGAAAACTTAAATCAAAAACAAGTTGCGCTTAAAACAAGTACACGTTCATTGGCTCGTTTA
GAAAATGCTCAACCAACATACGAAAAAGCAGTGAACGAGTTAAACAAGAGTTAAACAAGGTTGCAAGTGGTCCAAGCAAAAGAAGCC
TATGAAAATTCTCTGAAATCATTGGAGGAACTCAAAGAACAACAAGCCGTTGCTACACTTGCTTATACACAAGCAAGAA
GACCTTTCTAATGCGAAGTTAGAGCTACAGCAGTACCAAGGCATATTAAGAGACTTAGAGGCACAACAAGCCGAACAGCGC
CGACAAGAAGCGTTGCAAGAGCAAGTAGCAAAAGAGCAACAACGCCTTGAACGAGAAGCAAAGCAACAACAAATGTTAGTA
GCAAGTGCTACTTCAGCAGATAAAACACCTGGTCTCCAACAGTTATCTTTTTTCTAAACAAAAAGAACAGCCAAAAGCACAA

ACACTAACACATTCAGAACCTCGTAAGACGAAACAAGTAGCAAAAGCCCAAGAATCCTTACCACATACAGGAGAACAAAAA
AGTATCTGGTTGACTATTTTTGGATTATTCATGGCAGTAGGTGCGATCAGTTTCAAGAACAAAAGACGGAAAAATAGT

EFC0004
Seq ID 411
GTGTTGTTAGTATTTGCAAGTTGTGGAGTAAATAAATCAACAGAAGATTCAAAAAAAACAAATGAAACTAAGGTAGAACAA
ATAGCAACTTTATCGGCAGGTACACCAGTTCAAAGTTTAGATCCAGCAACAGCAGTGGATCAAACAAGTGTTACATTGTTG
GCCAACGTAATGGAAGGGTTGTACCGATTAGATGAAAAAAATCAACCACAACCAGCAATTGCAGCTGGGCAACCAAAAATA
TCCAATGGTGGTAAAACATATACAATTGTGATTCGAGACGGTGCAAAATGGGCCGATGGAACAGATATTACTGCTGATGAT
TTTGTAACGGCATGGCAAAGAGTTCTTGATCCTAAAACTGCCTCGCCAAACGTAGAACTGTTTGCTGCTATAAAAAATGCA
AAAGAGATTAGTATCGGAAAGCAACAAAAAGAAACACTTGGTGTAAAATCTAAGGGAAACAAAACAATAGAAATCGAATTA
GAAGAACCTACACCTTATTTTACAGATTTGCTAGCATTGACTGCTTATTTTCCAGTCCAACAAAACGCAGTAAAAGAGTAC
GGTAAAGAATATGGAACAACAAAAGAAACATTGTAACGAATGGAGCATTTACATTAACTGATTTAAACGGTGTTGGTATT
TCAGATAAATGGACGATTGCTAAAAAATCCAAAATACTGGGATAAGAAGCATGTCGCCATGGAAAAAATCAAATTTCAAGTT
GTAAAAGACATTAATACAGGTATTAATCTCTATAATGATGGTCAACTAGACGATGCGCCTGTAGCTGGGGAATATAGTAAG
CAATTAGAAAAATAACAAAGATTTCATACGAGAATTATCTGCTACAACTATGTTTTTAGAGGTAAATCAACGAAATAAAAAA
AGTATAAACTTCAAATAAACATGCTCGACAAGCGATTAATTTTTGCGGATTGATCGTGAGGCTATATCAAACAAGATATTAACA
AATGGAAGTATACCAGCTAAAGGTGTAGTACCGAGTAAACTAGTATATAATCCAAAAACAGGAAAAGATTTTACTAATTCA
AGTTTAGTGTTTTTTGGATAAGAGTAAAGCTAAAGACAGCTGGGAAAAAGCTAAAAAAGAATTAAAAGGTACTGATTTATCT
ATCGATATTATGGTCAATGAAGAAGATTTATCAAAAAAACTAGGTGAATATCTTCAAAATGAATTGCAAGATACTTTGGAT
GGACTTAAAGTTTCAGTTACTGCAGTACCTGCCACGTTACAAACTGAACGATTAAATTCTGGGAACTTTATGATTGCATTG
AGTGGTTGGCAAGCAGACTTTGCCGATCCAGTTAGTTTTTTAGAGAGTAAAAGTTCATTAAATCATGGTGGG
TATGCAAACGAAGAGTATGATAAATTGTTGAAAAATAATTCTTCAAAAAGATTGCAAGAATTGAAGGACGCCGAGAAATTA
ATTTTAGAAGATGCGGGTGTAATACCGTTACTTCAAATTGGAAATGCTAAATTAAGAAATCAAAAAATTTCTGAAATGAAA
GTACACTCTATTGGTGCTAAGTATGATTATAAAACCATGGAGATAAAA

EFC0012
Seq ID 412
ATGAAAAAAATTATTTTATCAAGCTTGTTTAGTGCAGTACTAGTATTCGGTGGCGGAAGTATAACAGCATTCGCTGACGAT
TTAGGACCAACAGATCCAGCAACTCCACCAATTACCGAACCAACTGATTCTAGTGAACCTACGAATCCTACTGAGCCGGTG
GATCCTGCAGAACCGCCAGTAATACCAACTGATCCAACAGAACCAAGCAAGCCAACCGAGCCTACAACACCGAGTGAGCCA
GAAAAGCCAACAGAACCAACAACGCCAATTGATCCTGGAACGCCGGTTGAACCGACTGAACCAAGCGAGCCAACAGAACCT
AGTCAACCAACCGAGCCTACAACACCAAGCGAACCAGAGAAAAACCTGTTACTCCAGAACAACCGAAAGAACCAACTCAACCA
GTGATTCCAGAAAAAACCAGCAGAACCAGAAACACCAAAAACTCCTGAACAGCCCACTAAACCAATAGACGTAGTCGTTACA
CCTAGTGGAGAAATTGATAAAACGAATCAATCGGCAGGAACCAACCAAGTATTCCTATTGAAACAAGCAACTTAGCGGAG
GTAACACATGTACCAAGTGAAACTACTCCAATTACAACAGAAGCTGGGGAAGAAATTGTAGCAGTAGATAAAGGTGTTCCG
TTAACCAAAACACCAGAAGGATTAAAAACCAATTAGCAGCTCGTATAAGGTTTTTACCTAGCGGAAACGTTGAGGTAAAAGCA
AGTGATGGAAAAATGAAAGTATTGCCACATACAGGAGAGAAATTCACACTCCTTTTCTCTGTATTGGGAAGCTTCTTTGTA
TTAATTTCAGGATTCTTTTTCTTTAAAAAGAATAAGAAAAAAGCT

EFC0021
Seq ID 413
ATGGTAGAAAGTAAGCGTACTTGGAAAGAAAAAGTACATGTTGAATCTCAAGTAGAAGAGCGTCTTAGAAAAAAATTTGACC
TTAAAATCAAGAGAAGCGACGAAGTTGCAAAAAAAAGCATTAAAGACGACTAAAAAGGAGTTTAAACTTTCTAAGAAAAAT
TACAAGCAATCACTAAAAAAAACATAAACTTACCTTGAAAACAGGGAAGGAACCATTAAGTATTAACCATTTACTGGATAAA
AAGAAATTTGTAGAGCAAAAGAAACGCAAAAAAGCACAGAAAAATTGCTTACAAACGAACGAAAAAAGTAGATGAAACGCGT
GTCGTTAACCAAGTGAAAAGGGAAACCAAGGAAGGGTTGAAAAGAGAAGCGACGCAAAAAGTACGAACCACGTTAACGCAG
GAAGATACACTGAATGAAGCAATGACCCTTTATGAAAAAACAACAAGCTAAATTTAACATGCGTACAGCTTTGAAAACA
GGAAAAACGGTTAAGAATCTTTCGGTGAAAACTGCAAAAGATACGTACGGATTGGGGAATCGCTTATTTAACTTTTCAAGA
GGACGTGGGTTTCAGCGCACACCAAAAGACTTTACGTTAAAAAAAACAACTTATGAAACAACTACGGAATCGCGCTATGCGC
TTTAAAGCAGCAAAAGAAGCGAAGAAAGCCGAACAAGGTATTGGCTTGATTCGTTCTTTTTTTAATGGTCAGAAGACATTA
GGGAAAGTCGCTGCATTGATTCTAAAAAATCCAATTTCTTGGGTTGTTTTATTGGTTCTTTTTTCTTGTTTTTTTATTGAGT
GGCGTTGCAAGTAGTACACAAAAACCAGCGATTGTTCAAGAAGAGGAAGATTTAACCGCTTCATGGACGTATTTTACTAAA
TTAGATGCGCAACATACAGACGATAACAACCTGTTTTATAGCAATATCGATGATGTTCTTTTCTACATGAATTATCGGTAC
GATGATTTTAAATTATTAGATATGGATTCCACTGGTACAAAGAATTTTGAAACGATTCTATCTGAACTTTGGACGGCACTA
AACGGAAAGAAACCTGATTATCAGTTGAAAACAATGCAAAGTTTAGAGACAGATAAGAAATCGAGTTATTTTATCGAAGAG
GAACAAGCAAAGCATTATCAAGAGATAAAAAAAGAGTTAGGCTACCAAACATTGGATGATCTCCTTTCTTTTCCAGTAAAA
ACTGACGCTTTAATTGTCAATAAGCGTTATGGCTATGACAAATCAAAAGAGAAGCTCACCTTATATAAAGGCATTGATGTA
TTGATTGAAGATAATCAGCCGTTTCATTCACCAATGAACGGCCAAATCGTTTCAGTGCCTGATACAGAAACGCTGGTTATC
GAAAAAGAGAAAGTAGCTAGGTTAACGATTCGTGGTGTGATACCCTTCGACTAACAAAAGGTATGGATGTTGAAGAAGGA
ACATTTCTTGGGAATACAAAGAATTCAACAGTTACGTTCCAATATGAAAAATATAAAAAAGAAACGAAAGACTGGTTTTTTT
GTAAATCCAGCCTTTTATTTTCCACGTGTAACCTACACGCAAACGACTCTTTTAGGAAGTGCAGAATTTTCTCCTGGAGCA
AGTGTCGAAAAAAGAGCGCAAGCTGTTTATGATTATTTATCAAAAAAGGGGATACACCAAAGAGGGAATTAGCGCGATTTTA
GGCAATTTTTTCGGTAGAAAGTGGCATTAATCCAAAACGTGCAGAAGGGGATTATTTGAATCCGCCTGTAGGTGCGCATGGG
AATTCATGGGACGAACCGAGTTGGCTTGCTATGGGAGGACCGCAAATTTATGGAGGACGTTTTCCAAATATTCTCCATCGA
GGACTAGGTTTAGGTCAGTGGACGGATACAGCTGATGGTGGTCGTAAGCGCATACTTTGTTATTAGACTATGCCAAAGGTAAA
AATAAAAAATGGTACGATCTGCAATTGCAACTGGACTTTATTTTTGACGGCGATGCACCAGGTTCACGAACAGCAGCAGAC
AATGTAGCAAGAAGTAAGGTGGCTGCAACTATTCCAGAACTTACTACTTATTATCTAACTGTTTGGGAAGGAAATCCCGGA
GACAAGTTAGGAGAACGAATTCAAGCTGCACAAAATTGGTTTACATTCTTTTCTCGATCGGGTACACCAATCGGTGGTTCT
GGGAGTGAGGTCTTTGCGCAGTATAAAGAAAAAAATGCAACCATTACCCACAGACAGAGAAACAAAAGCAGGACAAGGCTGG
CCAGGCAATGCGTATGCGCCAGGGAATTGTACTTGGTATGTGTTTAATCGTTTTGCACAGTTAGGAAAAAGCATTTACCCT

ACTATGGGAAACGCCAATCAATGGGTGCATAACTACAGTCAAACGCCCGGAGCGACACTTGAATCGGCCCCTAAAAAAGGG
GATGCTGTTATTTTTACAAATGGTGTAGCTGGTTCTAGTACGCAATATGGACACGTGGCTTATGTGGAGCATGTAAACTCT
GATGGTAGTTTTGTCATTTCTGAAATGAATGTCAGTGGTGAGTATTCTATGAATTGGCGTGTCTTAAAAAAAGAAGCAGGC
GAATACTTTATGCGCCTAAAC

EFC0053
Seq ID 414
ATGACCCAAGTACATTTTACACTGAAAAGCGAAGAGATTCAAAGCATTATTGAATATTCTGTAAAGGATGACGTTTCTAAA
AATATTTTAACAACGGTATTTAATCAACTAATGGAAAATCAACGAACAGAATATATTCAAGCAAAAGAATATGAACGAACA
GAAAACCGACAAAGTCAACGAAATGGCTATTATGAGCGCAGCTTTACGACACGTGTAGGCACGCTAGAATTAAAAGTACCC
AGAACACGTGATGGCCATTTTTCACCCACAGTGTTTGAACGTTATCAACGAAACGAAAAAGCCCTCATGGCTTCAATGTTG
GAAATGTATGTATCAGGCGTTTCAACTCGTAAAGTATCAAAAATTGTGGAAGAACTTTGTGGTAAATCCGTCTCTAAGTCC
TTCGTTTCTAGCTTAACAGAACAGCTAGAACCTATGGTTAACGAGTGGCAGAATCGTTTATTATCAGAAAAAAATTATCCT
TACTTAATGACCGATGTACTCTATATAAAAGTACGAGAAGAAATCGAGTACTCTCAAAAAGCTGTCATATAGCGATTGGA
ATAACCAAAGATGGCGACCGTGAAATTATCGGCTTCATGATTCAAAGTGGCGAAAGCGAAGAGACCTGGACAACATTTTTT
GAATACCTAAAAGAACGCGGTTTACAAGGTACGGAACTCGTTATTTCTGATGCGCACAAAGGATTAGTCTCTGCCATTAGA
AAATCCTTCACCAACGTAAGTTGGCAAAGATGCCAAGTTCACTTCCTAAGAAATATCTTTACCACCATTCCTAAAAAAAAAT
TCAAAATCTTTCAGAGAAGCTGTTAAAGGAATTTTTTAAGTTCACAGATATTAACTTAGCGCGTGAGGCTAAAAATCGATTG
ATTCATGATTATATCGATCAACCAAAATATTCAAAAGCTTGCGCATCATTGGATGATGGATTCGAAGACGCCTTTCAATAT
ACCGTACAAGGAAATTCCCACAATCGACTAAAGAGTACCAATCTAATTGAACGACTGAATCAAGAAGTACGCAGAAGAGAA
AAGATTATTCGCATCTTCCCCAATCAAACATCAGCCAATCGCTTAATTGGAGCCGTTCTTATGGACCTACATGATGAATGG
ATTTATTCTTCAAGAAAATACATCAATTTTGATAAG

ARFC0021.1
Seq ID 415
ccaagtgaaaagggaaaccaaggaagggttgaaaagagaagcgacgcaaaaagtacgaaccacgttaacgcaggaagatac
actgaa

ARFC0021.2
Seq ID 416
gagaacgaattcaagctgcacaaaattggtttacattcttttctcgatcgggtacaccaatcggtggttctgggagtgagg
tctttgcgcagtataaagaaaaaatgcaaccattacccacagacagagaaacaaaagcaggacaaggctggccaggcaatg
cgtatgcgccagggaattgtacttggtatgtgtttaatcgttttgcacagt

ARF0324
Seq ID 417
ctcaataaaagggcttgcgaaagtagaagctgtttacgatcaagaaatccaagaggtaagagaaacgtacaatgcgattat
tgcaaatgtagaaagagtgcctgc

ARF1627
Seq ID 418
gtgcctacattgcttacaaaccaactgtcggtatgccagaagcacggcggacagttattttcaaatattcacaaaggtggga
caccggctgttgaagcaggggcctacattgccgaaattattaaaaaaatgt

ARF1650
Seq ID 419
aagcggctgcaaaaaaagctgaaaaagtctatctcgcagccgatccggaccgagaaggtgaagcaattgcttggcatt

CRF0097
Seq ID 420
atggcattgttacctaatgtcacaagtaaaccagcccacggtgttgccattcctgtcacaggtacttttaagccaaagaag
gtcactacaatgccactaaccattcctgaaaataaaattcgttgcg

CRF0257
Seq ID 421
agttgcgcattttgccacgtttctttcgttggcaaatattttcgtagtccagctactggtacgaccacgattcctgcattc
gctgtcatcaggaaatttaacgcactcactcgatctgcgcgcgcttctggtgaagaaaaggccatttcggcggataacact
tcatctactggaaataga

CRF0635
Seq ID 422
ggaagcatcgtagataattcaccagatatcaatgcgattggaccagcaaagaataagaccgcgaaaattgtg

CRF1152
Seq ID 423
agccaggggcaatgttatggccaattaatgctgcttcaatacaaattgttccagaaccacaaactggatcataaaaaggac
gatctttgcgccaattggtcaacatcactaaagcagctgccatgttttcct

CRF1720
Seq ID 424

Tcattcgtccagtcctcctactgttcttcgctacgtgcgttttccggtaaaattaaattcaaaacaattcctaagacagtt
gctaatgccattgaagacaattcaaaagttcctactttgaagactaagccaccaatgccgatgactaaaattactgaagca
atcaataaattcttttttcttatcaaaattaatcttattatcgattaagattttcaaaccacttgcggcaatcacaccgaac
aagacgaaactaattcctgaaataactggtcctgggatacttagaatcaaggcacttaattttccaacaaagcctaaagct
acagcgaacaccgcggctccaccaatcacgaagacactgtgtactcttgtaattgctagcacgccgatatttttcacca

EF0008
Seq ID 425

MINNVVLVGRLTKDPDLRYTASGSAVATFTLAVNRNFTNQNGDREADFINCVIWRKPAETMANYARKGTLLGVVGRIQTRN
YENQQGQRVYVTEVVCENFQLLESRSASEQRGTGGGSFNNNENGYQSQNRSFGNNNASSGFNNNNNSFNPSSSQSQNNNGM
PDFDKDSDPFGGSGSSIDISDDDLPF

EF0028
Seq ID 426

MAVKKKKATFWEFFQGLGKTFMLPVALLAFMGILLGLGSSFSSESMIETIPFLGKPAVKIIFQFMSTIGGFAFAYLPVMFA
MAIPLGLVRKEKGIAAFSGFVGYTVMNLAINFYLVQTNRLVDPEQLREAGQGMVFGIQTIEMGVLGGIIAGLIVYKLHNRF
YTVQLPDSFAFFSGARFVPIITSLVMAFVGLVIPLVWPLFALMIMAIGQLIQRSGIFGPFLFGSGERLLLPPGLHHILVSM
IRFTEAGGSAVVAGKEVFGALNIFYAELQNNLPISPSATAFLSQGKMPTFIFGLPAASLAMYHTAAPANRHKIKGLLLSGV
IATAITGITEPIBFLFLFISPLLWLFHVIMTGLGFMVMALLGVVIGNTDGGLLDFVIFGLLQGTYTKWWWVLIVGAIWFVV
YYFVFKTVIVTFDLKTPGRDKVLDETEYTDQEVQYKKTGGYDAPGILAALGGQENIQAIDNCITRLRLVLADANKVDDDKL
KQLGALGVVHLDAQNVQVIIGTKVTTVRNQLEMILG

EF0146
Seq ID 427

MKKIASAGLSILVATGVAGIGGNEVQAAEQAQPKTPENSSTEQPAVKATETTEQAITEKQQQVTEKQAIVDQKQQVADTAK
KEKDTIDQSVKDQQAVVDQNKGALDQSQQAVTDQQAVVDEAKKVVDEATPSAIEKAKNQVATDTQAVDDQQKVVDQAQADV
NQQQAVVEEKAKETNAAKVQNDKDQQAVTAAKQEQAKLEELAKNAEVEKAKAEKEQAAKEAELANKQKEEAKAKDQKTKDD
QAVADQQTVVTTSQEKVADAKADTAAKQADLTAKENALKDKQTATKQVQNTLDKSKEELKGHKGINLPANFTPDYYKKLSE
QEKQAMEKEALALNKVFPENQADVAKATEMINVKNPTGKQKQQMSDYVVGLINDVREKLGLQKLKISNQAMRFAWDVAKYD
NPKEFDHDVNAINRAAKENGFKEYPGQNFYENLSMGYFETINGTISQLEFEKAARKTIADMLFDDESSAYSHIDSLLKGDT
TNMAVSISGDLNDISAKIHIISYNQSKLVEANTYEEGTAPVFKSKETLQKEVATNQEKLATAQQAESDAQQAKSASQQVLN
TAKTTQATAEKELSVHKATLASLQAVATKSTTNYEEKVRQTATINELIQNRAAVLEKAKTNVSEAQA
IEQTSAKVLEEKQAAQKAEENTLNSLKEVLNLAKENLNQKQVAFKTSTRSLSRLENAQPTYEKALNELNKAEAAAVQAQEA
YENSLKSLEELKEQQAVATLAYTQAQEDLSNAKLELQQYQGVLRELEAQQAEQQRQEALQEQVAKEQQRLEREAKQSQTLV
ASATSADKTPGLQQLSFSKQKEQPKAQALTHSESRKTKQVAKAPDSLPHTGEKNNKWLAIAGLIFALLGAAGIISFISRNE
KKVKNIFKIK

EF0153
Seq ID 428

MKKMIIALFSTSLLAGGSSVSAYAQESEGNLGETTGSVLPDEPNVPTDPITPSEPEQPTEPSTPEQPSEPSTPTEPSEPS
KPTDPSLPDEPSVPTEPTTPSKPEQPTEPTTPSVPEQPTEPSVPEKPVEPNKPTEPEKPVPVVPEKPVVPQQPEQPTDVVV
KPNGEIATGESTQQPTVPIETNNLSEVTHVPTVTTPIETASGEAIVAVDKGVPLTQTADGLKPIKSEYKVLPSGNVQVKSA
DGKMKVLPYTGEKMGIIGSIAGVCLTVLSGILIYKKRKV

EF0394
Seq ID 429

VKKRLFASVLLCSLTLSAIATPSIALADNVDKKIEEKNQEISSLKAKQGDLASQVSSLEAEVSSVFDESMALREQKQTLKA
KSEQLQQEITNLNQRIEKRNEAIKNQARDVQVNGQSTTMLDAVLDADSVADAISRVQAVSTIVSANNDLMQQQKEDKQAVV
DKKAENEKKVKQLEATEAELETKRQDLLSKQSELNVMKASLALEQSSAESSKAGLEKQKAAAEAEQARLAAEQKAAAEKAK
QAAAKPAKAEVKAEAPVASSSTTEAQAPASSSSATESSTQQTTETTTPSTDNSATENTGSSSSEQPVQPTTPSDNGNNGGQ
TGGGTVTPTPEPTPAPSADPTINALNVLRQSLGLRPVVWDAGLAASATARAAQVEAGGIPNDHWSRGDEVIAIMWAPGNSV
IMAWYNETNMVTASGSGHRDWEINPGITRVGFGYSGSTIVGHSA

EF0443
Seq ID 430

MKSIKTILLGTTLAAGLGLFLGTDANAESLYTVKAGDTLSTISHQFAGDNSLIQKIASDNKLPNLDLIFEGEQLVIRSEKE
VANTPAPAVEVAPVQQVVEQPVAQPVQQEVQQPVAQEVAQPAAPAASSDAKEWIAQRESSGSYDATNGQYIGRYQLSASYL
NGDYSPANQERVADQYVAGRYGSWDAAKSFWLANGWY

EF0568
Seq ID 431

MKKIYQWAVGQSFKKLDPRQQVKNPVMFVVYLGALITTILCFYPMGIPLWFNISITIFLWLTLLFANFAEAVAEGRGKAQA
DSLKQAKKEVMTYKINSLEDIKEENFIELQSSDLKRNDLVYVRAGEQIPADGDVIEGAASVDESAITGESAPVIRESGGDR
SAVTGGTTVVSDYLVIRVTSENGQSFLDKMIAMVEGTQRKKTPNEIGLQIFLITLTIIFLTVSITLVPFTDFSSQLSGKGE
ALSLVIVIALLICLAPTTIGALISSIGIAGMSRLTKENVIAMSGRAIEAAGDVDVLLLDKTGTITLGNRRASDFLPVHGVS
EEQLADAAQLSSLADETAEGRSIVILAKERFNLREREFQQSEVKFIDFSAKTRMSGIDYRGDVIRKGAADTMKKYVQSKGE
DYPSECDKIVDKIARAGGTPLVVIKNNRVMGVVYLKDIVKNGVKEKFADMRKMGIKTIMITGDNPLTAAAIAAEAGVDDFL
AEATPENKMNLIRBYQEKGHLVAMTGDGTNDAPALAQADVAMAMNTGTQAAKEAGNMIDLDSSPTKLLQVVQIGKQLLMTR

GALTTFSIANDIAKYFAVIPVLFYSIYPQLDRLNIMGLGSPLTAILSAVIYNAVVIVALIPLALKGVRYQEKPASQILSHN
LLIYGLGGIIAPFIFIKIIDLILSLIIL

EP0591
Seq ID 432
MFFTGKERNLKFAIKKKHGGGGAASVIIGCIIFGSIAFGGTTLAYADEVHNSINQDIQDSGSTIIGENDSSTKSAEYKMIH
EIDGTKISNGENSKETTTSSGTILAEEAIESSNQKNSKTSEVEQDLHKDVSGSESVKQVETSDSIKKSEESAVKTLNLDDS
QENTNSITTKAENDALSTVNDEKVLNESDSIIKSIPSETENVDNNSKSDNRDLVENSVETTKEIMKEADKTEIAKRNDNIS
VLNAGTESFNDSDKIRRSKRDNESPVPSLSIQMPVTKEIYENDSTIKGKSEPGVTIKIYKNGIELGSIKADDDGNFEYPLK
SQAIKSDNYSFVAEKDGAKSIESSVNVVENKEIVGNQRDAYTPYGFGIYNRGLVTTDDGKKHYTVDMEGNLASEAYFDLAG
GTMYYSIDEKFAPYVEKIVLDGKTIISKNPYELPNKTNIWSSGILSTDKRGGLVRAALVGRTNGTIDIYFKDDTPQEVLNS
EIPPFQVWARFKEKKEGVKNDLISDFDLNILLQNNKLILSKDTFFKDEDIVNNNGPKISSNFNYTNNTIDINYGVNTGNYGL
TIGKRSPYNLHINTGKELSDLVDTIKIDNKNYBFEKLPDGSLVIKDIYKKGILGGALLNRQEITISLGLIAKKNLGDLIDF
KQELLKLDVSIRDNTTFKLVKGEATEAIRKIDSSFDTFKQELTKWIEKRDSNPQEPITINSGDNTSIVDPKEMVEFAVKNP
TTENINKARSLVSKMWLWDKNKYNPVLNAAEEVNRYLSTTLNGSSNDDGLILNNTHAATFVKLIDTDKDGILDRYEANMGI
GTNYTNADSDGDGKSDGFEIFNDTDPLVSPYDWFDKNGEKIKIVTTDTDTISGRIGNNNYDTENVYPRTVQLIKVTDNGEN
LISEISSSEDRKGTFEFTGLSGKLSKGDRLVVKIITNEVQREIDRQRNVVQVGYDNPEVSEEVIVQGAQVTVSFDFNYNQT
DPDFDRQPVTQVVELEKGSRLLNSYPTPERKGYKFIGWNTDKLGVGQTVNDGSSFNEDITVFAQWEKEPNLAGEVHAPQGP
IEEGKSIVLNNVVQSNKSGSVIVTGELPKGLFGLSINETGNLIGTPLINDWIDGENSREVKIPVTISNGDEKVMVEVPLTI
LHDTRAADAVKAAEDAGKAGADKKAEVETDGLVTPEEKAAVDGLNDTTTAKKEDASKLVDALPEGPVKDSLKDRLDKVTTS
EVTVNDADSNGKADDVDLAEKAAADAVKAAEDAGKAGADKKAEVETDGLVTPEEKAAVDGLNDTTTAKKEDASKLVDALPE
GPVKDSLKDRLDKVTTSEVTVNDADSNGKADDVDLAEKAAADAVKAAEDAGKAGADKKAEVETDGLVTPEEKAAVDGLNDT
TTAKKEDASKLVDALPEGPVKDSLKDRLDKVTTSEVTVNDADSNGKADDVDLAEKAAADAVKAAEDAGKAGADKKAEVETD
GLVTPEEKAAVDGLNDTTTAKKEDASKLVDALPEGPVKDSLKDRLDKVTTSEVTVNDADSNGKADDVDLAEKAAADAVKAA
EDAGKAGADKKAEVETDGLVTPEEKAAVMA

EP0592
Seq ID 433
LPEGPVKDSLKDRLDKVTTSEVTVNDADSNGKADDVDLAEKAAADAVKAAEDAXKAGADKKAEVETDGLVTPEEKAAVDGL
NDTTTAKKEDASKLVDALPEGPVKDSLKDRLDKVTTSEVTVNDADSNGKADDVDLAEKAAADAVKAAEDAGKAGADKKAEV
ETDGLVTPEEKAAVDGLNDTTTAKKEDASKLVDALPEGPVKDSLKDRLDKVTTSEVTVNDADSNGKADDVDLAEKAAADAV
KAAEDAGKAGADKKAEVETDGLVTPEEKAAVDGLNDTTTAKKEDASKLVDALPEGPVKDSLKDRLDKVTTSEVTVNDADSN
GKADDVDLAEKAAADAVKAAEDAGKAGADKKAEVETDGLVTPEEKAAVDGLNDTTTAKKEDASKLVDALPEGPVKDSLKDR
LDKVTTSEVTVNDADSNGKADDVDLAEKAAADAVKAAEDAGKAGADKKAEVETDGLVTPEEKAAVDGLLEIKQSSFMPFEN
LFSTTNDYSQFPKTGEKSDSILTIYGGLLFLSSIGLLGIKKRKNNTN

EP0658
Seq ID 434
MKRKINYTHIAEPTVPENVFPTYLTNETMMARKQKVLQRMETEKFDQLVFYADKEHGSNFEYLTGFIPRFEEGLLVLNKDG
AATLILGNENLKLCQHARISADLIHYLAFSLPNQPMAGEQKLSQIFETLLDETAQKIGIVGWKMFTTQQQEPATMFDVPHF
IVEALKKALPKEARLLNGTHLLIGAKGARATNNANELAHYEYGANLASRSMLKALNAIEIGQRETDIGALLNDEGQTPTVV
TIAATGQRFEYANMYPTAKEIQLGDALSLTTGYKGGLSSRTGFVIENEQQLPEAQRDYLERVAKPYFQAVVHWLETIRIGL
LGREMYQAIKEQLPKEIYHWHLNPGHLVSDDEWMSSPIYPDSAIRLESGMLFQVDIIPSVPGYTGVSAEECVALADETLQK
EIQQTYPDMWQRIATRKAYLKETLKLDLPSEVLPMSNLVGYLRPFYLAKDKALCVEKPALK

EP0727
Seq ID 435
MKKARVIYNPTSGKELIKKNLADILSILEECGYEASAFATTPEENSARNEAHRAARAGFDLLVAAGGDGTINEVVNGIAPL
KRRPKMAIIPAGTTNDYARALKIPRDNIVKAAEVIKKNQTVKMDIGQAGKNYFINIAAGGHLTELTYEVPSELKSIFGYLA
YLAKGAEMLPRVKPIKMRMTYDEGVYEGNASMPFLGLTNSVGGFEQIVPDAKLDDGKFSLIIVKTANIFEILHLVALMLNG
GKHVEDHRLIYTKTSYLHAETLEKNNKMMINLDGEYGGDAPMTFKNMHQHIEIFANGDALPSNAIMGSVLTGSDEIVVESE
DEEEAYNEASKEFVKEVERLTDEDIDGDGKIAEKEKH

EP0775
Seq ID 436
MTNTVKVKDDSLADCKRILEGQATFPVQAGETEPVDLVVVEDASGSFSDNFPHVRQAIDEVVQGLSDQDRVMLASYRGGKQ
PMFPDGKTKINSADYDMNVRVNTQLTYDKSQFVSGFGDVRTYGGTPTAPGLKLALDTYNQTHGDLTNRKTYFLLVTDGVAN
TRLDGYLHKTNTNDSINEYPDPRHPLQVSVEYSNDYQGAAAEVLALNQEITNQGYEMINAYWESVESLSSVNSYFDKYKTE
VGPFVKQELQQGSSTPEDFITSQSIDDFTTQLKQIVKDRLAQSTPATASLTIANQFDIQSATATDDAGNDVPVQINGQTIS
ATSTEGYVGNITIHYEVKENTAIDAATLVSSGTMNQGTIAKEFPEATIPKNDNAHACDVTPEDPTITKDIENQEHLDLTNR
EDSFDWHVKTAFGNETSTWTQASMVDDINKVLDIIDVKVTDENGKDVTANGTVTQENNKVTFEMNKQADSYDYLSGHTYTM
TITTKIKTDATDEELAPYIEQGGIPNQADLNFGNEGDVLHSNKPTVTPPPVDPNIAKDVEGQEHLDLTNRDQEFKWNVKTA
FGNETSTWTQASMVDDINKVLDITDVKVTDENGKDVTANGKVTQENNKVTFEMNKQADSYDYLSGHTYTMTITTKIKASAT
DEELAPYIEQGGIPNQADLNFGNEGDVLHSNKPTVTPPAPTPEDPTITKDIEGQEHLDLTNRDQEFKWNVKTAFGNETSTW
TQASMVDDINKVLDITDVKVTDENGKDVTDNGIVTQENNKVTFTMNKKDDSYSYLAGHTYTMTITTKIKTDATDEELAPYI
EQGGIPNQADLNFGNEGDVLHSNKPTVTPPAPTPEDPKKPEPKQPLKPKKPLTPTNHQAPTNPVNFGKSASKGIHLPMTNT
TVNPLYMIAGLIVLIVAISFGITKNKKRKN

EP0779
Seq ID 437

MSAWQTFKNGTRQPFFKDILQYLWLFFTLNVLLLLLVGGAFSWATSNALKTQGIPYLSFNNLNLLLEKPLALVLLILLLLLFL
GAVFYQPTFLLLGIFQIRQDHRFHFKGVTKASFKVLKKQGARSWLFFFGYFVVIVPFGNLIFQSNLLTKFVIPDFIVEFLS
QRIPYLVGLLAPGLLVWYLAIRFIYTLPLMILERKKAGEAVKASWSMTNKRLWFIIRNIAFVTIAVFVSTYVIYVLLYLLQ
LKLDTLSDTISLLGGIILNLTVVQFLQFVSNAWLSVLLINFLYTQLNVQAETTTKVAFDKETKRNKLVTIGMGLGLFTIFGG
YIIFNAVYLTGLLESKPLIISHRGVTNSNGVQNTIPAMERTIKFKPDYIEIDVQETKDHQFVVMHDANLQELAGVDGTPQE
FTLAELTKMTVKENGQEAPIASFDDYLAKANQAKQKLLVEIKTSKQDSQGALSNFIEKYERPLIKNNHQVQSLDYNVIKAF
KKAKSKVKVSFILPYNFTFPETQADLYTMEATTLNDTFILKADQQKKAVYAWTVNDSEVLSKMLFMDVAGVITDDLELVNE
EVNDFEKNPSYADRILHYIFMLPSVASQ

EF1091
Seq ID 438
MITDENDKTNINIELNLLNQTEQPLQREIQLKNAQFMDTAVIEKDGYSYQVTNGTLYLTLDAQVKKPVQLSLAVEQSSLQT
AQPPKLLYENNEYDVSVTSEKITVEDSAKESTEPEKITVPENTKETNKNDSAPEKTEQPTATEEVTNPFAEARMAPATLRA
NLALPLIAPQYTTDNSGTYPTANWQPTGNQNVLNHQGNKDGSAQWDGQTSWNGDPTNRTNSYIEYGGTGDQADYAIRKYAR
ETTTPGLFDVYLNVRGNVQKEITPLDLVLVVDWSGSMNENNRIGEVQKGVNRFVDTLADSGITNNINMGYVGYSSDGYNNN
AIQMGPFDTVKNPIKNITPSSTRGGTFTQKALRDAGDMLATPNGHKKVIVLLTDGVPTFSYKVSRVQTEADGRFYGTQFTN
RQDQPGSTSYISGSYNAPDQNNINKRINSTPIATIGEAMVLKQRGIEIHGLGIQLQSDPRANLSKQQVEDKMREMVSADEN
GDLYYESADYAPDISDYLAKKAVQISGTVVNGKVVDPIAEPFKYEPNTLSMKSVGPVQVQTLPEVSLTGATINSNEIYLGK
GQEIQIHYQVRIQTESENFKPDFWYQMNGRTTFQPLATAPEKVDFGVPSGKAPGVKLNVKKIWEEYDQDPTSRPDNVIYEI
SRKQVTDTANWQTGYIKLSKPENDTSNSWERKNVTQLSKTADESYQEVLGLPQYNNQGQAFNYQTTRELAVPGYSQEKIDD
TTWKNTKQFKPLDLKVIKNSSSGEKNLVGAVFELSGKNVQTTLVDNKDGSYSLPKDVRLQKGERYTLTEVKAPAGHELGKK
TTWQIEVSEQGKVSIDGQEVTTTNQVIPLEIENKFSSLPIRIRKYTMQNGKQVNLAEATFALQRKNAQGSYQTVATQKTDT
TGLSYFKISEPGEYRMVEQSGPLGYDTLAGNYEFTVDKYGKIHYAGKNIEENAPEWTLTHQNNLKPFDLTVNKKADNQTPL
KGAKFRLTGPDTDIELPKDGKETDTFVFENLKPGKYVLTETFTPEGYQGLKEPIELIIREDGSVTIDGEKVADVLISGEKN
NQITLDVTNQAKVPLPETGGIGRLWFYLIAISTFVIAGVYLFIRRPEGSV

EF1323
Seq ID 439
MGILLPIIIAVLVLLMLLIVFVSKYQTAKPDEALIISGSYLGSKNVHVDEGGNKIKIVRGGGAFVLPVFQRSNRISLLSSK
LDVSTPEVYTEQGVPVMCDGTSIIKIGSSVEEIATAAEQFLGKTTEELENEAREVLEGHLRSILGSMTVEEIYQNRDFSQS
VQEVASVDLAKMGLVIVSFTIKEVRDKNGYLDSLGKPRIAQVKRDADIAEAEALKETRIKKAEAEKESQQAELQRQTEIAE
ASKEKELKLALYKQEQDIAKAKADQAYNLESARAQQHVVEQEMEVKVVERQKQIELEKKEITRREKQYDSEVKKKADADRY
AREQEALAQKAREVAEAEAERFKVEALAEAEANKTRLTGQAQAEAILARGAAEAEAKQKIADAFKEYGEAAVLSMVMEMLP
QLMKEAAQPLGNIDKISVVDTGAGGENSGANRITNYATNLLAGTQETLKETTGLDVKELIENFSKKGTSNSVNYHATEGSE
KE

EF1355
Seq ID 440
MAYQFKLPDIGEGIAEGEIVKWFVKPGDTINEDDTLLEVQNDKSVEEIPSPVTGTVKNIVVPEGTVANVGDVLIEIDAPGH
EDNDAAPAAPAQEQTPAQPAAVPTTEAAGGFFQFKLPDIGEGIAEGEIVKWFVKAGDTINEDDSLLEVQNDKSVEEIPSPV
TGTVKNIVVPEGTVANVGDVLVEIDAPGHNSAAPSVAAPATDAPKAEASAPAASTGVVAAADPNKRVLAMPSVRQYAREKD
VDITQVTATGKGGRVIKADIDAFVSGGSQAAPATEAAATEAAPKAEAAAPKAAPKAFTSDLGEMFTREKMTPTRKAIAKAM
VNSKHTAPHVTLHDEVEVSKLWDHRKKFKDVAAANGTKLTFLPYVVKALTSTVQKFPILNASIDDAAQEIVYKNYFNIGIA
TDTDHGLYVPNVKNANTKSMFAIADEINEKAALAIEGKLTAQDMRDGTITISNIGSVGGGWFTPVINYPEVAILGVGTIAQ
EPVVNADGEIVVGRMMKLSLSFDHRIVDGATAQKAMNNIKRLLADPELLLMEG

EF1699
Seq ID 441
MEYTIKKMASLSGVSARTLRYYDEIGLLQPARINSSGYRIYGQAEVNRLQQILFYRELDLKLDEIKEILEQPDFNVEQALY
EHQQKLLEKRNEIDRLLASVQQTLHHYKGEINMSDQQKFEAFKQQKVQENEEKYGKEIREKYGNETIEQANKKYLNLTEKD
MQAMQNVEKDLFSKLAMYQKSPKLTSQLAQEIFQLHKDWLMYSWSSYSPEAHKGLGLMYVGDERFTSYYEQHGAGFAEALN
AIIQNYA

EF1744
Seq ID 442
MAKKGGFFLGAVIGGTAAAVAALLLAPKSGKELRDDLSNQTDDLKNKAQDYTDYAVQKGTELTEIAKQKAGVLSDQASDLA
GSVKEKTKDSLDKAQGVSGDMLDNFKKQTGDLSDQFKKAADDAQDHAEDLGEIAEDAAEDIYIDVKDSAAAAKETVSAGVD
EAKETTKDVPEKAAEAKEDVKDAAKDVKKEFKG

EF1752
Seq ID 443
MKRKLTKSPNNVVLTGSLAGIADWLGIDPTIIRVVYVLLSFFSAGPPGILLYIALAVLIPSGRTGSDRGYGHQNPYNRNVR
NENPYAANKKQRKEAEKIDDDEWSDF

EF1753
Seq ID 444
MKERERVLELVKKGILTSEEALILLENMATEKDEKQIEKAAEKVDTQNIGTTNKEDQVADLMNALEKGESEGPTVDSFEEN
TQDSAEKDRENLERILDELATKANRASAELDEVNAEIAGIKEEIKEVAEEIGTLDTKEELDALTEDEQVQRKDLHVLLAQL
EEKLATQSTEKTALEEELKNIRKEQWKGQWNDTKEKVSSQFSEEWKDQATDTFNQVGGKVAEVGGQVGEFLKKTFNSFSDT
MNDNVEWKDIKMKVPGVATTKFEHEFNYPNPQASLIDVKVANGTVVFKTWDQEDVKVEAKIKLYGKMAGDSPMEAFLERSD
IDVDDETISFQVPNKRVKADLTFYLPKRTYDEVSVKLLNGNVLEELTAKDVYTKSTNGTITFKKIDATMLEIEGVNGEIK

VLEGTILDNIIETVNGDVSISAAPESLSVSLINGDIRITAKEKTLRRVEASSANGNIKLALPNDLGVEGQVKTNLGSINSR
LTDIEVVREKKDRGNQQLHFRRVLEESMAQINASTTTGSIFLKDTDK

EF1791
Seq ID 445
MKKYLKITMVCILLVGFLAGCTNKNENKKKQKNTKEAVQLMSPSELTTLNTSVLLDFPDAIVQTAAPEGLYSLDEQDQLVP
AVAKALPMISEDGKTYTISLRKEAVWSNDDPVTAHDFRYAWKKMIDPKNGFVYSFLIVETIQNGAEISAGKLAPNELGVTA
VDDYTLKVTLKEPKPYFTSLLAFPTFFPQNQKVVEQFGADYGTASDKVVYNGPFVVKDWQQTKMDWQLAKNNRYWDHQNVR
SDIINYTVIKETSTALNLFEDGQLDVATLSGELAQQNKNNTLYHSYPTATMNYLRLNQKRKGQATPLANENLRKALALGID
KENLVNNIIADGSKALHGAITEGFVANPTTGLDFRQEAGNLMVYNKEKAQSYWKKAQAELGEKVNVELMVTDDGSYKKIGE
SLQGSLQELFPGLTIELTALPTEAALNFGRESDYDLFLIYWTPDYQDPISTLMTLYKGNDRNYQNPVYDKLLDEAATTYAL
EPEKRWATLIAAEKEVIETTAGMIPLSQNEQTVLQNDKVKGLNFHTFGAPLTLKNVYKEK

EF1800
Seq ID 446
MKHGKIKRFSTLTLLASATILVPLSTSAEETTNSSTETSSSMVEPTATEEKLWQSDFPGGKTGEWQDVIGKTNRELAGESL
AISRDAAAGNNAVSLNLDSPKLADGEVETKFKYTAGSGRTGVIIRGNTKDSWVFVGYNANGKWLVESPNSWNDSISGPTLN
EDTNYLLKVRYVGEKITIWLNTTLIYEGEPVLANGDKIPTEAGHVGVRLWYDKKIVNYDYFKNGPVDSIPEIVPEVTQIAP
VKVFTKIGVAPKLPKQVKVTYNTGKEANEAVRWNEIDPDAYKEPGTFEVDGTLENTNIKAKASIVVAKDNEAEKGDKISSA
DLTAVVDPQFPRIIRYEDPQSNQVIFNGQHEKIDQVMIDGKAYKATAEKQKSEANQAVYNAVPEIGLRFTTTLTVSEGQE
LAMKLSDIREEGTKIHTISIPNQGLISVNSTDEGATFAGVVMNTGTNANNGNKNGDTIQDLTTTSQEETKKYMYGFLNTAN
YAASFWTNAYGDGSVDGSDNNRIHKQTKEAATGFVTTLSSGAWTYRPFDAPEDYTTGETPEVKVKFSKDSNDDNRVDWQDA
AIGFRSIMNNPMGAEKVPELVNQRIPFNFASQATNPFLVTLDESKRIYNLTDGLGQMNLLKGYQNEGHDSAHPDYGAIGQR
PGGEQALNQLIDEGHKLNAVFGVHINDTESYPEAKGFNEELVDPTKRGWDWLDPSYFIKQRPDTLSGRRYERFKELKQKAP
NLDYIYVDVWGNQGESGWASRQLSKEINSLGWFTTNEFPNALEYDSVWNHWSAEKDYGGTTTKGFNSTIVRFIRNHQKDTW
IISDNPLLGGAEFEAYEGWVGKTNFNTYRQKTFAINVPTKFLQHYQITNWETTTAADGQIYGTIKLANGAEKVTVTQADAN
SPRSITLNETEVLKGDAYLLPWNVNGQDKLYHWNPKGGTSTWSLDKKMQGKTNLHLYELTDQGRIDKGAIATTNNQVTIQA
EANTPYVIAEPDSIEPMTFGTGTPFKDPGFNEANTLKNNWKVPRGDGEVKKDANGDYVFSSEKERTEIKQDINLPKPGKYS
LYLNTETHDRKATVTVKIGGKKYTRTVNNSVAQNYIQADINHTSRKNPQYMQNMRIDFEIPDNAKKGSVTLAVDKGNSVTK
FDDLRIVERQTDIMNPDKQTVIKQDFEDTQAVGLYPFVKGSAGGVEDPRIHLSERNEPYTQYGWNGNLVSDVLEGNWSLKA
HKQGAGLMLQTIPQNIKFEPNKKYTVQFDYQTDGENVFTAGTINGELKNNNDFKPVGELTSTAADGQTKHYEAEIIGDASG
NTTFGIFTTGADKDFIMDNFTVTVESKK

EF1818
Seq ID 447
LMKGNKILYILGTGIFVGSSCLFSSLFVAAEEQVYSESEVSTVLSKLEKEAISEAAAEQYTVVDRKEDAWGMKHLKLEKQT
EGVTVDSDNVIIHLDRNGAVTSVTGNFVDQVVKIQSVDAIGEEGVKKIIASDNPETKDLVFLAIDKRVNNEGQLFYKVRVT
SSPTGDPVSLVYKVNATDGTIMEKQDLTEEVGSEVTLKNSFQVAFNVPVEKSNTGIALHGTDNTGVYHAVVDGKNNYSIIQ
APSLVALNQNAVDAYTHGKFVKTYYEDHFQRHSIDDRGMPILSVVDEQHPDAYDNAFWDGKAMRYGETSTPTGKTYASSLD
VVGHEMTHGVTEHTAGLEYLGQSGALNESYSDLMGYIISGASNPEIGADTQSVDRKTGIRNLQTPSKHGQPETMAQYDDRA
RYKGTPYYDQGGVHYNSGIINRIGYTIIQNLGIEKAQTIFYSSLVNYLTPKAQFSDARDAMLAAAKVQYGDEAASVVSAAF
NSAGIGAKEDIQVNQPSESVLVNE

EF1850
Seq ID 448
VQQNNNEISTQHQFDAYCKKVLRNEAKSIRKRNDKIKENEEPLNDLNEGKYSQTHLDEQDVYFLFGMEILISDQKLSAAID
QLSDTRKKIVLLYYFAGFNDTEIGKIFNMSTSGIWYQRTKAVEQLKMEYGLW

EF1877
Seq ID 449
MKKKLFRIVGIVFISILTLIVLLSLVGTVAEATGLVDDTVESGNLYSKYSLNNYQLDFFVDSSWDWLPWNWGDGLGKSVMY
GLYAITNFIWTVSLYLSNATGYVVQEAYKLDFISDTAESIGKNIQTLADITENGLQTSGFYFGFLLLMILALGVYVAYTGL
LKRETTKAVRAVINFVMIFLLSGSFIAYAPTYITKINDFSSDVSEAALTLGTEIVVPNSESQGEKDSVDLIRDSLFSIQVQQ
PWLLLQFDDSNIEEIGEDRVNKILSVSPDENKGKDREEAVKAEIEDNDNANLSITKTMNRLGIVVFLVLFNIGISPFVFLL
TGIMLFSQILFIIFAMFLPISFLLSMLPTYESLGKMAIIRLFNTIMMRAGVTLVITTAFSISTMFFNISATYPFFMVAFLQ
IVTFAGIYFKLGDIMSMFNLQSNDSQSHGRRVMRRPQMLMNRKLRQLNRNVGRTLAFGGAAAVGNKLAKEQSKPKFKPAGS
SLRKNSRLPNDREVSSDSAKENPISNNKKQSRMNLTGRKIGKVLDTQALVKDKAKQVKDQVRNTPTNLKYILHKGLEKTKK
APKEFKRGLVQEKADREKLRDKQRQRRDEKMDEKRKNLGEVTDRHGKRRSNVSIKEDPKPQKDRILKNELPKRIVMVKPNS
EVRRRLVKQNLVARENNQLLSKKNSFQQVKQRSTLPKRANQKVQRIMKARPKPKSGDKK

EF2174
Seq ID 450
MKKIISGMLICTVLLNSPSVIASGEELVKTETTGETGLVTQATSETTTNSTEDTSSVTEENTSERDSSTSSTKEESLDSST
NSTTEEHSSVSETNTTDSKTEASQSSEVEKKTIDQDEADYQEAAREGTNHKKGTYAMKNGLSSRVARATVANVYANDPNLP
GKNFIDVSSWNGDISVAEYQKIKSYGVTGVSVKLTEGTWYVNPYAAGQIRNAKAAGLKVSAYHYSMYVSAATAQDEARYFA
QAAANSGLDKNTIMFNDAEDPTLTNNGRNAHANSVAFNQQLKALGYKNDALYVGKWWLTNGYIDTSAFGRDRVNVAQYPYT
PDSSMQWNNDHGAWQWSSQMYFPGLANYEGRPFDISMTYSNFLNMGNSSGPDLSKYYTTNPGRVIMKNDDTFYQDVAFRTP
GWRVKKNTLVTIKGIEYSSAGIPRLVTDQGYLTANKDYVLAAQSNIDLYFTTNPKKVRLKSDDYFYADPEFKQRLSKVSKG
TIVEVEDLAYTQSGIFRLKTAKGYLTANKNIVEQYKGLEDTYYTSNPGQVITRNEDTYYKDVEFKTKANKVSAGSVLKVTA
IEKTKSGIPRLKMANGYYFTANKNYVVATGSWIANYHTVNPGQIIMKNSDNFYGDSDFLYKGAAVSKGSLVPVVGIEYREN
QVPRLITQNGYLTANKSYAQKVVPNIKDYLYDYPEYVVMKTNDYYYQDVNFSKKGEFVSKDTLIKVLSVDYTENGLPRLKT

AKGYITANKSYVTKLVSNSDNYFTENPHQIIMRVSDKLYTDVEFRNGSRTLSSGTVVPVKGIEYSSKGVPRLKTEGGYLTA
NKNYVTAAGNTNNNYPITNPKKVKLLIDDCFYNNTEFTKKGQAVKKNTIVEVEAIEYTNNGIPRLKTKQGYLTANKWYVAK
VG

EF2224
Seq ID 451
MNKAVKNFVSYLMITMLFILNLLPMMNAFAQEVTSDAEKTVEKDGLKVIGKIEDTSSQEDIKTVTYEVTNTRDVPIKDLIL
KQKNTNDSPIKFVLDTLSEERGPTSLEEQAKVETNEKDQTTDIKLLNLQPNSTRKITINGQITTKASNKLLVSVLIEDNEK
GTLVIDLPSKDILADKESVSKEKQETSETKVENQANETASSTNEMTATTSNETKPEAGKAIESIQETALTQATESPEQPPL
KAQPTGPLVPPTPGRGFNTPIYQSVHKGELFSTGNTNLKIANENTAAAQTFLNTRGASSGYAINNFPLEFADVDNDPNTYN
SSRAYIDLNGAKEIAWAGLFWSASRYKGPAYGTNLSDEEISAPVQFTTPNGTVQRVSPQRYHRIDQDATNPGQRFGYNNTG
FSNYADVTSILQGDKSATGSYTLADIPMTSSLNGQYQYYNFSGWSLFVVTKDQASKSRAFSIYYGARGNAAGTNNEFTMSN
FLTAKQGNLDPIVTWFTVQGDKYWTGDNAQIKNSAGTWVNISNTLNPVNNAMNATVTDNDEHMVDKYPGKFAPDHPNFLDI
DIDRMAIPEGVLNAGONQINFRTTSSGDDYSTNAIGFAVNAETPEFEIKKEIVEPKETYKVGETITYRVSLKNTKADSEAI
NSVSKDALDGRLNYLPGSLKIISGPNSGEKTDASGDDQAEYDETNKQIIVRVGNGATATQGGSYKADTAETIYEFKARINE
RAKANELVPNSATVEAVDILTSAKVNETSNIVEAKIADEQVTGKLTATKTVNNAKPKLGEEIEYTISFRNTIENGILNKVV
ITDQLPKGLTYVKDSLTSVGDEPKPTSLKETNGTITAEYPSITDMKERSIRFKVIVNEEAKAGETILNKAKVDDTVNPPEE
PEVPVVPETNAGKLAATKTVNNAKPKLGETIEYTISFRNTIENGVLNKVVITDQLPKGLTYVKDSLTSVGDEPKPTSLKEI
NGTITAEYPSITDTKERSIRFKVIVNEEAKAGETILNKAKVDDTVNPPEEPEVPVVPETNVGKLTATKTVNNAKPKLGEEI
EYTISFRNTIENGVLNKVVITDQLPKGLTYVKDSLTSVGDEPKPTSLKEINGTITAEYPSITDTKERSIRFKVIVNEEAKA
GETILNKAKVDDTVNPPEEPEVPVVPEAKEGKLTATKTVNNAKPKLGEAIEYTISFRNTIENGVLNKVVITDQLPKGLTYV
KDSLTSVGDEPKPTSLTEANGTITAEYPSITDMKERSIRFKVIVNEEAKAGETILNKAKVDDTVNPPEEPEVPVVPEAKEG
KLAATKTVNNAKPKLGETIEYTISFRNTIENGVLNKVVITDQLPKGLTYVKDSLTSVGDEPKPTSLKEANGTITAEYPSIT
DTKERSIRFKVIVNDEAKAGETILNKAKVGDGINPPEEPEVPITPEEPAKNKKETNKVVTDQNKPTKNSKNEIAINKKETS
KSSYLPKTGEKVQKIFAYLGVGLILIVLILYVIKRNKEKEE

EF2318
Seq ID 452
VSRERFQAREKKVQKMGRDGLVEQNRATGEEKRVSQRTADTAFDRARPTEREDIRRPAARGADTGKKRKQPRTAPEMAAEQ
TEPLLTPEYLADSTAEPPPDAPASMRGAVDMPLLDAPIAEEPTPAPVRKQPQKRGKKGKKKQSTKFTEDAARPGSEEKTAM
RGVEGAPAPKRRTSEGGSRLQFDKEEEPADPAADRAAIRKRQTAKLAEHAAKPEDKLKSDGNPKAGETRLKFEDTPKDTDI
SGDTAPDMPTRQQKKYNKAACRTERADRRVEQAQMKLPAKRRLHMDHQPDRASGKIKRRLRFEDEILPEYQKPSLSSRAGN
AAKTAAVLKLHGKLREYERDNVALESAHKIELVAEQGAGRVLRWERNRRRSKPYRTLRKAQQKAARAHTDLAWQTALRDNP
ELQRKNALAKNVQKQKIKRKYAQAAHEAKQSAKFTQNVLTATGKIARAIAQYAAAHKAVFLAVAMLALVVMFFATGLTSCT
AMLSGFQSSYISASYMANEQEICQSDLYYTEKETDLQIDIDKTEENYPGYDEYRYNIGEISHNPYELLGYLSTAFNAFTPA
EVQPEIDRIFSRQYTLTREVIVETRYDDDGDPYDWYVLQTTLAVRPLSSVLQSSLTPGEQTDRYGVYMQTYGNRQAFGNPF
GFSWLGYVSSGYGWRVHPVNGEKSLHRGIDLAVAQGTPILAAQDGRVVSAGDAGSYGLCVVIEDDKGYQSRYAHCSSLNVS
AGQEVKRGDVIAAVGSTGNSTGPHLHLEVMLNGEYLNPYFFVDNGDDGTGAIPGTPGGPAIPDYSGEPMGDGSFEAMLREA
EKYLGYPYVWGGSHPSTSFDCSGYVSWVINQSGVGSVGRQTAQGLYNLCTPVSAVNAQPGDLIFFTGTYSSPGPVSHVGIY
VGGGRFIHCGDPISYANTGSPYWSAHLYGYGRIP

EF2704
Seq ID 453
MTNKEPWEAWSPAKVSSFQEDFLAWYEREKRNLPWRANTDAYRIWISEIMLQQTRVDTVIDYFYRFMEWFPTIQDLAEAPD
DKLLKAWEGLGYYSRARNLKVAAQQIVSEFGGKMPDTIEDIRSLKGIGPYTAGAIGSIAFNLREPAIDGNVMRVVSRLFEI
DADIAKASSRKVFEAAMLKIIDRERPGDFNQALMDLGSAVCTPTSPKCESCPLQQYCAAYQADKMTAYPVKSKKVKPKDVY
YVGTIIENKKQEFLLEQRPETGLLANMWLFPIEEISKKQFQQLQKLAQPAETEKQLTLELEPVTEPLVAEEPVNFFTDYET
VVWQKRTLGEVVHIFSHLKWHILVFYGRNTGELATLESQRWVAAQQFSDYVFPKPQQKMVELFKKEHEKK

EF2713
Seq ID 454
VEKSTVTSLTQETSATTINASTDSTALTAESEELPSLRQTLLNYVGMYGLTETLINRLSDDELDYAKKVSFHFVNQDISGT
ARMITKLYGEKPIPEDSYSTDYSTLTIDDLKNYLPQIRLSLIYVYDLNSDVVNNLSDQTLVDLINQVKVDYANQNYPSDVR
GDYGLAAMADKIKANDYTSINQSAESVSSDTTNTESTLQTTTSSSKKATTSSSTEHKKGIFPSTGEKKSVLFTIIGIILLS
LVSIFIIKNKKK

EF2802
Seq ID 455
MKKKILVGALVALFFMPLNVFAAKGDQGVDWAIYQGEQGRFGYAHDKFAIAQIGGYNASGIYEQYTYKTQVASAIAQGKRA
HTYIWYDTWGNMDIAKTTMDYFLPRIQTPKNSIVALDFEHGALASVPDGYGGYVSSDAEKAANTETILYGMRRIKQAGYTP
MYYSYKPFTLNHVNYQQIIKEFPNSLWIAAYPIDGVSPYPLYAYFPSMDGIGIWQFTSAYIAGGLDGNVDLTGITDSGYTD
TNKPETDTPATDAGEEIEKIPNSDVKVGDTVKVKFSVDANATGEAIPDWVKGNSYKVQEATGSRVLLEGILSWISKGDIEL
LPDATVVPDKQPETTHIVQHGETLSSIAYQYGTDYQTLASLNGLANPNLIYPGQVLKVNGSATSNVYTVQLGDNLSSIAAK
LGTTYQTLAALNELANPNLIYPGQTLNY

EF2813
Seq ID 456
VSKQESDVVLNFKMNGEINYSRTIKDINKEMNLAATEYKNQVSAMDKNATQTEKLTATKKKLEKQLSLAEQRTKLLREEYE
KSVKETGEYSEQSQKLYKRLLESETGENKLRSALQSTNEALKEQGNLSIKTAEKLAKIEKAGDKIKSVGQKLSVGLTAPIM
GIGAASIAAFKELDECLDNITTATGATGSQLESLQASFKTVAGQIPADMQDISTGIGEVNTQFGLMDKQLEDTTGRMLKFS
EINGSDVSQSTINAKKSMDLFRLSIEDLPMILDSVSKTSQDTGVGVDQLFDAVNRGAPQLKAMGLGFSESTTLIGQMEKAG

IDSAGTLGYLAKASVVYAKDNKTMQEGLSGTIESIKGATTEQEKLTIASEVFGTKAASKMVEAIDSGALSMDGLADSAKNA
AGTVDQTFSDILDPIDQAKLAQNQFKIAMGELGEQVQIALLPAFQAATDAIKKVSEWFGSLTDSQKQTILKIAGVVAAIGP
VLVVLGTLASSISSLIPVIAFIASPIGLVIAAVAAWVAAIVVAYNKIGWFRDFINTSPKVIKDIVVGVFNVLKDTTKSTFD
FITGFIGGAMDGAARIIGDYVNEIKRIFGGIVDFVTGVFTGDWSRAWQGVVDIFGGIFEGIAAVAKAPINAMITLINGPIG
GLNNIKIPKWVPGIGGKGFHIGKIPYLAEGGTILNGQAIVGEAGPELLTAKNGKTTVTPLSPEEKARGIGGALKGGNTIEQ
HVHIGQVDANNPSELDRMNRKLYKASAQAFYDLGGVPT

EF2820
Seq ID 457
MNKELLRQLQARHEKRLSDLQGKIESGEVREADLDSVNEEIDGLIDELKAIKAELGDDNSESGDGEGDDGTAKSDNTDDES
KEDREKDTNGNNDDKNEENRGGMISQEQRDGLLRTIHEGMEARNVMSNEQREKQIRKAFADFVIGNISESEARALGIETGN
GSVTVPEVIASEVISYAQEENLLRKYGTVIRTAGDVKYPILVKKAEANVNKKERTTDITETAIQFDEILLDPAEFDALATV
TKKLLKMSGVPVEDIVVEELKKAYVRKEINYMFNGDDAGNENPGALAKKAVAFEKPVDLTAAGAGQKLYDALIEFKNTPVT
EVMKKGRFIINRAALTAIEKMKTDDGFPLLRPFTQAEGGIGYQLVGYPVDWTDAADKKGEPDTPVLYFGDFSAPKIQEVIG
ALEIQKLVEKFSGKNQVGFQIYNLLDGQLVYSPFEPAVYRYEITKPVGG

EF3082
Seq ID 458
MKKKFLAMMAVSMMGLLMLSACQTNKKTADSATTETTAKTEVTVKDTNGQLTVPKNPKKVVVFDNGSLDTMDALGVGDRVV
GAPTKNIPAYLKKYQKVESAGGIKEPDLEKINQLKPDLIISGRQQDYQEQLKAIAPTIYLAVDAKNPWASTKQNIETLGT
IFDKEEVAKEKITGLEKEIADVKKQAEASANNALVVLVNEGQLSAYGKGSRFGLIHDTFGFKAADDKIEASTHGQSVSYEY
VLEKNPGILFVVDRTKAIGGDDSKDNVAANELIQKTDAGKNDKVIMLQPDVWYLSGGGLESMHLMIEDVKKGLE

EF3256
Seq ID 459
MKVNKFVKGFAAIALSSLVLAACGADKKDNTTNSSSAASSETKKSTESSAPAKKVAGGDLKDGTYKLEEKNEKNGYRAVFE
MTVKDGKITESKYDNINADGKSKTEDTKYEESMKAKSGVGPKEYIKQLNDSFVKAQSASGVEVVTGATHSSESFQNYAQQL
IQAAQAGNTDTIEIDNGATLKDGTYSLKEKNDSNGYHTTFSMTVKDGKVTESNYDNVNADGKSKKDDTEYESKMKDVAGVG
PKEYIETLNKEFVKAMGEEDGSPAGVEVVTGATHSTHSFINYAQQLVNAAEKGDTTEIVVDNIVTK

EFA0021
Seq ID 460
MDFQNIQTVKELRYALKQYKQTLATTIFDDTQGYISVAFKKEVIEKDGKNVQFLCFYPNYEVALYPVETLLAYLEKLPEQA
IVTIKDHDGNYSSLFLTIDETIDQGNNYQWLVISDKNSYERFFFTAQESTQEQAFEKMLKELETDHDSTLEVVHNWLCRQT
DEALFLGILKEDRTLKGAMTYCIGKAQEQSKDQSSAMVADEVVFSWIKEYFLLEKLPETKAVGKVTTKKSETKPERKEVDE
QIDLFECV

EFA0044
Seq ID 461
MIILSGLFILGIGIVGGYQLATLPKLIEMKQHKAIQNHFNVKGNEYTYYQEDSENYILSLEDTEYRIKFSKNTPLKVVFTE
ILEPM

EFA0052
Seq ID 462
MKKIASTGLSILVATGVAGIGGNEVQAAEQAQPKTPENSSTEQPTVKATQTTEQAITEKQQQVTEKQAIVDQKQQVADTAK
KEKDAIDQSVKDQQAVVDQNKDALDQSQQAVTDQQAVVDEAKKVVDEATPSAIEKAKEQVATDTQAVDDQQKVVDQAQTDV
NQQQAVVEEKAKETNAAKVQNDKDQQAVTAAKQEQAKLEELAKNAEAEKAKAEKEQAAKEAELANKQKEEAKAKDQKTKDD
QAVADQQSVVTTSQEKVADAKADTAAKQADLTSKENALKDKQAATKQAQNTLDSSKEELKGHKGLNLPANFTPDYYKKLTE
QEKQAMEKEALALNKVFPENQADAAKVTEMINVKNPTEKQKQQMSDYVVGLINDVREKLGLQKLKISNQAMKFAWDVAKYD
NPKEFDHDVNAINRAAKENGFKEFPGQNFYENLSMGRPTTQEGKVSMYDFEKAARNALVSMLMNDGHSGYSHLDSLLDANE
TNMAVSISGDLNDISAKIHIISYNQTKLVEANTYEEGTAPVFKSKETLQKEVATNQEKLATAQQAESDAQQARSASQQALN
TAKTTQATAEKELSVHKATLANLQAVATKSTTNYEEKVRQTAAAETNLQQTKDQLATINELIQNRAVVLEKAKTKVADAQA
IEQTSAKVLKEKQEAQKAEENTLNSLKEVLDLAKENLNQKQVALKTSTRSLARLENAQPTYEKAVNELNKAEAAVVQAKEA
YENSLKSLEELKEQQAVATLAYTQAQEDLSNAKLELQQYQGILRDLEAQQAEQRRQEALQEQVAKEQQRLEREAKQQQMLV
ASATSADKTPGLQQLSFSKQKEQPKAQTLTHSEPRKTKQVAKAQESLPHTGEQKSIWLTIFGLFMAVGAISFKNKRRKNS

EFC0004
Seq ID 463
VLLVPASCGVNKSTEDSKKTNETKVEQIATLSAGTPVQSLDPATAVDQTSVTLLANVMEGLYRLDEKNQPQPAIAAGQPKI
SNGGKTYTIVIRDGAKWADGTDITADDFVTAWQRVLDPKTASPNVELFAAIKNAKEISIGKQQKETLGVKSKGNKTIEIEL
EEPTPYFTDLLALTAYFPVQQNAVKEYGKEYGTTKENIVTNGAFTLTDLNGVGISDKWTIAKNPKYWDKKHVAMEKIKFQV
VKDINTGINLYNDGQLDDAPVAGEYSKQLENNKDFIRELSATTMFLEVNQRNKKSITSNKHARQAINFAIDREAISNKILT
NGSIPAKGVVPSKLVYNPKTGKDFTNSSLVFLDKSKAKDSWEKAKKELKGTDLSIDIMVNEEDLSKKLGEYLQNELQDTLD
GLKVSVTAVPATLQTERLNSGNFMIALSGWQADPADPVSFLANFESKSSLNHGGYANEEYDKLLKNNSSKRLQELKDAEKL
ILEDAGVIPLLQIGNAKLRNQKISEMKVHSIGAKYDYKTMEIK

EFC0012
Seq ID 464
MKKIILSSLFSAVLVFGGGSITAFADDLGPTDPATPPITEPTDSSEPTNPTEPVDPAEPPVIPTDPTEPSKPTEPTTPSEP
EKPTEPTTPIDPGTPVEPTEPSEPTEPSQPTEPTTPSEPEKPVTPEQPKEPTQPVIPEKPAEPETPKTPEQPTKPIDVVVT

PSGEIDKTNQSAGTQPSIPIETSNLAEVTHVPSETTPITTEAGEEIVAVDKGVPLTKTPEGLKPISSSYKVLPSGNVEVKA
SDGKMKVLPHTGEKFTLLFSVLGSFFVLISGFFFFKKNKKKA

EFC0021
Seq ID 465
MVESKRTWKEKVHVESQVEERLRKNLTLKSREATKLQKKALKTTKKEFKLSKKNYKQSLKKHKLTLKTGKEPLSINHLLDK
KKFVEQKKRKKAQKIAYKRTKKVDETRVVNQVKRETKEGLKREATQKVRTTLTQEDTLNEAMTLYEKTQQAKFNMRTALKT
GKTVKNLSVKTAKDTYGLGNRLFNFSRGRGFQRTPKDFTLKKQLMKQLRNRAMRFKAAKEAKKAEQGIGLIRSFFNGQKTL
GKVAALILKNPISWVVLLVLFLVFLLSGVASSTQKPAIVQEEEDLTASWTYFTKLDAQHTDDNNLFYSNIDDVLFYMNYRY
DDFKLLDMDSTGTKNFETILSELWTALNGKKPDYQLKTMQSLETDKKSSYFIEEEQAKHYQEIKKELGYQTLDDLLSFPVK
TDALIVNKRYGYDKSKEKLTLYKGIDVLIEDNQPFHSPMNGQIVSVPDTETLVIEKEKVARLTIRGVNTLRLTKGMDVEEG
TFLGNTKNSTVTFQYEKYKKETKDWFFVNPAFYFPRVTYTQTTLLGSAEFSPGASVEKRAQAVYDYLSKEGYTKEGISAIL
GNFSVESGINPKRAEGDYLNPPVGAHGNSWDEPSWLAMGGPQIYGGRFPNILHRGLGLGQWTDTADGGRRHTLLLDYAKGK
NKKWYDLQLQLDFIFDGDAPGSRTAADNVARSKVAATIPELTTYYLTVWEGNPGDKLGERIQAAQNWFTFFSRSGTPIGGS
GSEVFAQYKEKMQPLPTDRETKAGQGWPGNAYAPGNCTWYVFNRFAQLGKSIYPTMGNANQWVHNYSQTPGATLESAPKKG
DAVIFTNGVAGSSTQYGHVAYVEHVNSDGSFVISEMNVSGEYSMNWRVLKKEAGEYFMRLN

EFC0053
Seq ID 466
MTQVHFTLKSEEIQSIIEYSVKDDVSKNILTTVFNQLMENQRTEYIQAKEYERTENRQSQRNGYYERSFTTRVGTLELKVP
RTRDGHFSPTVFERYQRNEKALMASMLEMYVSGVSTRKVSKIVEELCGKSVSKSFVSSLTEQLEPMVNEWQNRLLSEKNYP
YLMTDVLYIKVREENRVLSKSCHIAIGITKDGDREIIGFMIQSGESEETWTTFFEYLKERGLQGTELVISDAHKGLVSAIR
KSFTNVSWQRCQVHFLRNIFTTIPKKNSKSFREAVKGIFKFTDINLAREAKNRLIHDYIDQPKYSKACASLDDGFEDAFQY
TVQGNSHNRLKSTNLIERLNQEVRRREKIIRIFPNQTSANRLIGAVLMDLHDEWIYSSRKYINFDK

ARFC0021.1
Seq ID 467
psekgnqgrvekrsdakstnhvnagryte

ARFC0021.2
Seq ID 468
enefklhkiglhsfldrvhqsvvlgvrslrsikkkcnhypqtekqkqdkagqamrmrqgivlgmclivlhs

ARF0324
Seq ID 469
lnkracesrsclrsrnprgkrnvgcdyckcrksac

ARF1627
Seq ID 470
vptlltnqlsvcqkhggqlfqiftkvghrilkqgptlpkllkkc

ARF1650
Seq ID 471
krlqkklkksisqpirtekvkqllgi

CRF0097
Seq ID 472
mallpnvtskpahgvaipvtgtfkpkkvttmpltipenkfva

CRF0257
Seq ID 473
scafchvsfvgkyfrspatgtttipafavirkfnaltrsarasgeekaisadntsstgnr

CRF0635
Seq ID 474
gsivdnspdinaigpaknktakiv

CRF1152
Seq ID 475
sqgqcygqlmllqyklfqnhkldhkkddlcanwstslkqlpcfp

CRF1720
Seq ID 476
sfvqssycsslrafsgkikfktipktvanaiednskvptlktkppmpmtkiteainkffflsklillsikifkplaaitpn
ktklipeitgpgilrikalnfptkpkatantaappitktlctlviastpifsp

(51) International Patent Classification⁷: C07K 14/315, C12N 15/31, C12Q 1/68, G01N 33/569

(21) International Application Number:
PCT/EP2004/005664

(22) International Filing Date: 26 May 2004 (26.05.2004)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
03450137.9 30 May 2003 (30.05.2003) EP

(71) Applicant (for all designated States except US): INTER-CELL AG [AT/AT]; Campus Vienna Biocenter 6, A-1030 Wien (AT).

(72) Inventors; and
(75) Inventors/Applicants (for US only): MEINKE, Andreas [DE/AT]; Piettegasse 26/1, A-3013 Pressbaum (AT). NAGY, Eszter [HU/AT]; Taborstrasse 9, A-1020 Wien (AT). HANNER, Markus [AT/AT]; Jacquingasse 5/6, A-1030 Vienna (AT). GELBMANN, Dieter [AT/AT]; Ungergasse 5, A-7163 Andau (AT).

(74) Agent: SONN & PARTNER PATENTANWALTE; Riemergasse 14, A-1010 Wien (AT).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Declaration under Rule 4.17:
— of inventorship (Rule 4.17(iv)) for US only

Published:
— with international search report

(88) Date of publication of the international search report:
16 June 2005

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: ENTEROCOCCUS ANTIGENS

(57) Abstract: The present invention discloses isolated nucleic acid molecules encoding a hyperimmune serum reactive antigen or a fragment thereof as well as hyperimmune serum reactive antigens or fragments thereof from E.faecalis, methods for isolating such antigens and specific uses thereof.

INTERNATIONAL SEARCH REPORT

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7   C07K14/315   C12N15/31   C12Q1/68   G01N33/569

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7   C07K   C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the International search (name of data base and, where practical, search terms used)

EPO-Internal, Sequence Search, PAJ, WPI Data, BIOSIS

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DATABASE Geneseq 'Online! 19 June 2003 (2003-06-19), "Protein encoded by Prokaryotic essential gene #14577." XP002303524 retrieved from EBI accession no. GSN:ABU29050 Database accession no. ABU29050 The sequence is 100% identical with SEQ ID NO: 171 & WO 02/077183 A (MALONE CHERYL ; OHLSEN KARI L (US); WALL DANIEL (US); XU H HOWARD (US)) 3 October 2002 (2002-10-03) -/--	1,2, 5-28, 30-37

[X] Further documents are listed in the continuation of box C.   [X] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
2 November 2004	2 8. 01. 05

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL – 2280 HV Rijswijk Tel. (+31–70) 340–2040, Tx. 31 651 epo nl, Fax: (+31–70) 340–3016	Grosskopf, R

Form PCT/ISA/210 (second sheet) (January 2004)

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DATABASE UniProt 'Online! 1 June 2003 (2003-06-01), "PTS system, mannose-specific IIAB components. VLFLVDLWGG TPFNQANSLL EDHKDKWAIV AGMNLPMVIE AYASRFSMES AQEIATHILE" XP002303525 retrieved from EBI accession no. UNIPROT:Q839X9 Database accession no. Q839X9 The sequence is 100% identical with SEQ ID NO: 171	1,2, 5-28, 30-37
X	DATABASE Geneseq 'Online! 22 April 2004 (2004-04-22), "Enterococcus faecalis polypeptide #1547." XP002303526 retrieved from EBI accession no. GSN:ADH87067 Database accession no. ADH87067 100% identity with SEQ ID NO: 171 & US 6 617 156 B1 (DOUCETTE-STAMM LYNN A ET AL) 9 September 2003 (2003-09-09)	1,2, 5-28, 30-37
X	DATABASE Geneseq 'Online! 1 January 2004 (2004-01-01), "E. faecium protein sequence SEQ ID 6586." XP002303527 retrieved from EBI accession no. GSN:ADC96959 Database accession no. ADC96959 The sequence has 86% identity with SEQ ID NO: 171 & US 6 583 275 B1 (DOUCETTE-STAMM LYNN A ET AL) 24 June 2003 (2003-06-24)	1,2, 5-28, 30-37
X	DATABASE Geneseq 'Online! 18 December 2002 (2002-12-18), "Enterococcus faecalis contig sequence #464." XP002303528 retrieved from EBI accession no. GSN:ABS99196 Database accession no. ABS99196 The sequence is 100% identical with SEQ ID NO: 1	1,2,5-10

-/--

INTERNATIONAL SEARCH REPORT

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category°	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	DATABASE UniProt 'Online! 1 October 2001 (2001-10-01), "Mannose-specific phosphotransferase system component IIAB. VLFLVDLWGG TPFNQANSLF EEHKDKWAIV AGMNLPMVIE AYGARLSMES AHEIAASIIS" XP002303529 retrieved from EBI accession no. UNIPROT:Q97TN2 Database accession no. Q97TN2 _____	
A	HENICS T ET AL: "Small-fragment genomic libraries for the display of putative epitopes from clinically significant pathogens" BIOTECHNIQUES, EATON PUBLISHING, NATICK, US, vol. 35, no. 1, July 2003 (2003-07), pages 196-200,202,20, XP002293668 ISSN: 0736-6205 _____	
T	MEINKE ANDREAS ET AL: "Bacterial genomes pave the way to novel vaccines" CURRENT OPINION IN MICROBIOLOGY, vol. 7, no. 3, June 2004 (2004-06), pages 314-320, XP002303523 ISSN: 1369-5274 _____	

Box II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☒ Claims Nos.:   29
because they relate to parts of the International Application that do not comply with the prescribed requirements to such
an extent that no meaningful International Search can be carried out, specifically:

see FURTHER INFORMATION sheet PCT/ISA/210

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box III Observations where unity of Invention Is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all
searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment
of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report
covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is
restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

1, 2, 5-28, 30-37 (all partially)

Remark on Protest   ☐ The additional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM   PCT/ISA/ 210

This International Searching Authority found multiple (groups of)
inventions in this international application, as follows:

Invention 1:   Claims 1, 2, 5-28, 30-37 (all partially)

> Claims relating to a nucleotide sequence and the
> corresponding proteins sequence having SEQ ID NOs: 1 and
> 171, the use of said sequences and antibodies directed
> against said protein

Inventions 2 to 166: Claims 1, 2, 5-28, 30-37 (all partially)

> Claims relating to the nucleotide sequences and the
> corresponding proteins sequence having SEQ ID NOs: 2 to 170
> and 172 to 340 (with the exception of SEQ ID NOs: 19, 28,
> 78, 93 respectively 189, 198, 248 and 263), the use of said
> sequences and antibodies directed against said protein

Inventions 167-218: Claims 3, 5-28, 30-37 (all partially)

> Claims relating to a nucleotide sequences and the
> corresponding proteins sequence having SEQ ID NOs: 373 to
> 424 and 425 to 476 respectively , the use of said sequences
> and antibodies directed against said protein

---

4. claims: Inventins 219-220: Claims 4-28, 30-37 (all partially)

> Claims relating to a nucleotide sequences and the
> corresponding proteins sequence having SEQ ID NOs: 90 and
> 147 and 260 and 317 respectively, the use of said sequences
> and antibodies directed against said protein

FURTHER INFORMATION CONTINUED FROM   PCT/ISA/ 210

Continuation of Box II.2

Claims Nos.: 29

Claim 29 relates to an agonist which lacks any meaningful technical
characterisation. Therefore a search for accordingly defined compounds
was not possible

The applicant's attention is drawn to the fact that claims relating to
inventions in respect of which no international search report has been
established need not be the subject of an international preliminary
examination (Rule 66.1(e) PCT). The applicant is advised that the EPO
policy when acting as an International Preliminary Examining Authority is
normally not to carry out a preliminary examination on matter which has
not been searched. This is the case irrespective of whether or not the
claims are amended following receipt of the search report or during any
Chapter II procedure. If the application proceeds into the regional phase
before the EPO, the applicant is reminded that a search may be carried
out during examination before the EPO (see EPO Guideline C-VI, 8.5),
should the problems which led to the Article 17(2) declaration be
overcome.

Patent document cited in search report		Publication date	Patent family member(s)		Publication date
WO 02077183	A	03-10-2002	US	2002061569 A1	23-05-2002
			WO	02077183 A2	03-10-2002
			US	2004029129 A1	12-02-2004
US 6617156	B1	09-09-2003	NONE		
US 6583275	B1	24-06-2003	NONE		